09/869,877

PCT/DK 01/00292

Kongeriget Danmark

Patent application No.: PA 2000 00707

Date of filing: 28 April 2000

Applicant: Novozymes A/S
Krogshøjvej 36
DK-2880 Bagsværd

This is to certify the correctness of the following information:

The attached photocopy is a true copy of the following document:

- The specification, claims, abstract and sequence listing as filed with the application on the filing date indicated above.

By assignment dated 17 Nov 2000 and filed on 01 Dec 2000, the application has been assigned to Novozymes A/S

Patent- og
Varemærkestyrelsen
Erhvervsministeriet

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

TAASTRUP 21 May 2001

Lizzi Vester
Head of Section

PROTEIN VARIANTS HAVING MODIFIED IMMUNOGENICITY

Field of invention

The present invention relates to a method of selecting a protein variant having modified immunogenicity as compared to the parent protein, to the protein variant and use thereof, as well as to a method for producing said protein variant.

Background of the invention

An increasing number of proteins, including enzymes, are being produced industrially, for use in various industries, housekeeping and medicine. Being proteins they are likely to stimulate an immunological response in man and animals, including an allergic response.

Depending on the application, individuals get sensitised to the respective allergens by inhalation, direct contact with skin and eyes, or injection. The general mechanism behind an allergic response is divided in a sensitisation phase and a symptomatic phase. The sensitisation phase involves a first exposure of an individual to an allergen. This event activates specific T- and B-lymphocytes, and leads to the production of allergen specific IgE antibodies (in the present context the antibodies are denoted as usual, i.e. immunoglobulin E is IgE etc.). These IgE antibodies eventually facilitate allergen capturing and presentation to T-lymphocytes at the onset of the symptomatic phase. This phase is initiated by a second exposure to the same or a resembling antigen. The specific IgE antibodies bind to the specific IgE receptors on mast cells and basophils, among others, and capture at the same time the

allergen. The polyclonal nature of this process results in bridging and clustering of the IgE receptors, and subsequently in the activation of mast cells and basophils. This activation triggers the release of various chemical mediators involved in the early as well as late phase reactions of the symptomatic phase of allergy. Prevention of allergy in susceptible individuals is therefore a research area of great importance.

For certain forms of IgE-mediated allergies, a therapy exists, which comprises repeated administration of allergen preparations called 'allergen vaccines' (Int. Arch. Allergy Immunol., 1999, vol. 119, pp1-5). This leads to reduction of the allergic symptoms, possibly due to a redirection of the immune response away from the allergic (Th2) pathway and towards the immunoprotective (Th1) pathway (Int. Arch. Allergy Immunol., 1999, vol. 119, pp1-5).

Various attempts to reduce the immunogenicity of polypeptides and proteins have been conducted. It has been found that small changes in an epitope may affect the binding to an antibody. This may result in a reduced importance of such an epitope, maybe converting it from a high affinity to a low affinity epitope, or maybe even result in epitope loss, i.e. that the epitope cannot sufficiently bind an antibody to elicit an immunogenic response.

There is a need for methods to identify epitopes on proteins and alter these epitopes in order to modify the immunogenicity of proteins in a targeted manner. Such methods and kits for their execution can have at least four useful purposes:

1) reduce the allergenicity of a commercial protein using protein engineering.

2) reduce the potential of commercial proteins to cross-react with environmental allergens and hence cause allergic reactions in people sensitized to the environmental allergens (or vice versa).

3) improve the immunotherapeutic effect of allergen vaccines.

4) assist characterization of clinical allergies in order to select the appropriate treatment, including allergen vaccination.

In WO99/53038 (Genencor Int.) as well as in prior references (Kammerer et al, Clin. Exp. Allergy, 1997, vol. 27, pp 1016-1026; Sakakibara et al, J. Vet. Med. Sci., 1998; vol. 60, pp. 599-605), methods are described, which identify linear T-cell epitopes among a library of known peptide sequences, each representing part of the primary sequence of the protein of interest. Further, several similar techniques for localization of B-cell epitopes are disclosed by Walshet et al, J. Immunol. Methods, vol. 121, 1275-280, (1989), and by Schoofs et al. J. Immunol. vol. 140, 611-616, (1987). All of these methods, however, only leads to identification of linear epitopes, not to identification of 'structural' or 'discontinuous' epitopes, which are found on the 3-dimensional surface of protein molecules and which comprise amino acids from several discrete sites of the primary sequence of the protein. For several allergens, it has been realized that the dominant epitopes are of such discontinuous nature (Collins et al., Clin. Exp. All. 1996, vol. 26, pp. 36-42).

Slootstra et al; Molecular Diversity, 2, pp. 156-164, 1996 disclose the screening of a semi-random library of synthetic

peptides for their binding properties to three monoclonal antibodies by immobilizing the peptides on polyethylene pins and binding a dilution series of each antibody to the pins. This reference does not disclose any indication of how the antibody binding peptide sequences relate to any full protein antigens or allergens.

In WO92/10755 a method for modifying proteins to obtain less immunogenic variants is described. Randomly constructed protein variants, revealing a reduced binding of antibodies to the parent enzyme as compared to the parent enzyme itself, are selected for the measurement in animal models in terms of allergenicity. Finally, it is assessed whether reduction in immunogenicity is due to true elimination of an epitope or a reduction in affinity for antibodies. This method targets the identification of amino aicds that may be part of structural epitopes by using a complete protein for assessing antigen binding. The major drawbacks of this approach are the 'trial and error' character, which makes it a lengthy and expensive process, and the lack of general information on the epitope patterns. Without this information, the results obtained for one protein can not be applied on another protein.

WO 99/47680 (ALK-ABELLÓ) discloses the identification and modification of B-cell epitopes by protein engineering. However, the method is based on crystal structures of Fab-antigen complexes, and B-cell epitopes are defined as "a section of the surface of the antigen comprising 15-25 amino acid residues, which are within a distance from the atoms of the antibody enabling direct interaction" (p.3). This publication does

not show how one selects which Fab fragment to use (e.g. to target the most dominant allergy epitopes) or how one selects the substitutions to be made. Further, their method cannot be used in the absence of such crystallographic data for antigen-antibody complexes, which are very cumbersome, sometimes impossible, to obtain - especially since one would need a separate crystal structure for each epitope to be changed.

Hence, it is of interest to establish a general and efficient method to identify structural epitopes on the 3-dimensional surface of commercial and environmental allergens.

Summary of the invention

The present invention relates to a method of selecting a protein variant having modified immunogenicity as compared to a parent protein,

comprising the steps of:

a) obtaining antibody binding peptide sequences,

b) using the sequences to localise epitope sequences on the 3-dimensional structure of parent protein,

c) defining an epitope area including amino acids situated within 5 Å from the epitope amino acids constituting the epitope sequence,

d) changing one or more of the amino acids defining the epitope area of the parent protein by genetic engineering mutations of a DNA sequence encoding the parent protein,

e) introducing the mutated DNA sequence into a suitable host, culturing said host and expressing the protein variant, and

f) evaluating the immunogenicity of the protein variant using the parent protein as reference.

A second aspect of the present invention is a protein variant having modified immunogenicity as compared to its parent protein. The amino acid sequence of the protein variant differs from the amino acid sequence of the parent protein with respect to at least one epitope pattern of the parent protein, such that the immunogenicity of the protein variant is modified as compared with the immunogenicity of the parent protein.

A further aspect of the present invention is a composition comprising a protein variant as defined above, as well as the use of the composition for industrial application, such as the production of a formulation for personal care products (for example shampoo; soap; skin, hand and face lotions; skin, hand and face crèmes; hair dyes; toothpaste), food (for example in the baking industry), detergents and for the production of pharmaceuticals, e.g. vaccines.

Yet another aspect is a DNA molecule encoding a protein variant as defined above.

Further aspects are a vector comprising a DNA molecule as described above as well a host cell comprising said DNA molecule.

Another aspect is a method of producing a protein variant having modified immunogenicity as compared to the parent protein as defined above.

Definitions

Prior to a discussion of the detailed embodiments of the invention, a definition of specific terms related to the main aspects of the invention is provided.

In accordance with the present invention there may be employed conventional molecular biology, microbiology, and recombinant DNA techniques within the skill of the art. Such techniques are explained fully in the literature. See, e.g., Sambrook, Fritsch & Maniatis, Molecular Cloning: A Laboratory Manual, Second Edition (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (herein "Sambrook et al., 1989") DNA Cloning: A Practical Approach, Volumes I and II /D.N. Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed. 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds (1985)); Transcription And Translation (B.D. Hames & S.J. Higgins, eds. (1984)); Animal Cell Culture (R.I. Freshney, ed. (1986)); Immobilized Cells And Enzymes (IRL Press, (1986)); B. Perbal, A Practical Guide To Molecular Cloning (1984).

When applied to a protein, the term "isolated" indicates that the protein is found in a condition other than its native en-

vironment, such as apart from blood and animal tissue. In a preferred form, the isolated protein is substantially free of other proteins, particularly other proteins of animal origin. It is preferred to provide the proteins in a highly purified form, i.e., greater than 95% pure, more preferably greater than 99% pure. When applied to a polynucleotide molecule, the term ''isolated'' indicates that the molecule is removed from its natural genetic milieu, and is thus free of other extraneous or unwanted coding sequences, and is in a form suitable for use within genetically engineered protein production systems. Such isolated molecules are those that are separated from their natural environment and include cDNA and genomic clones. Isolated DNA molecules of the present invention are free of other genes with which they are ordinarily associated, and may include naturally occurring 5′ and 3′ untranslated regions such as promoters and terminators. The identification of associated regions will be evident to one of ordinary skill in the art (see for example, Dynan and Tijan, Nature 316: 774-78, 1985).

A ''polynucleotide'' is a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5′ to the 3′ end. Polynucleotides include RNA and DNA, and may be isolated from natural sources, synthesized in vitro, or prepared from a combination of natural and synthetic molecules.

A ''nucleic acid molecule'' refers to the phosphate ester polymeric form of ribonucleosides (adenosine, guanosine, uridine or cytidine; ''RNA molecules'') or deoxyribonucleosides (deoxyadenosine, deoxyguanosine, deoxythymidine, or deoxycytidine; ''DNA molecules'') in either single stranded form, or a double-stranded helix. Double stranded DNA-DNA, DNA-RNA and

RNA-RNA helices are possible. The term nucleic acid molecule, and in particular DNA or RNA molecule, refers only to the primary and secondary structure of the molecule, and does not limit it to any particular tertiary or quaternary forms. Thus, this term includes double-stranded DNA found, inter alia, in linear or circular DNA molecules (e.g., restriction fragments), plasmids, and chromosomes. In discussing the structure of particular double-stranded DNA molecules, sequences may be described herein according to the normal convention of giving only the sequence in the 5′ to 3′ direction along the nontranscribed strand of DNA (i.e., the strand having a sequence homologous to the mRNA). A ''recombinant DNA molecule'' is a DNA molecule that has undergone a molecular biological manipulation.

A DNA ''coding sequence'' is a double-stranded DNA sequence, which is transcribed and translated into a polypeptide in a cell in vitro or in vivo when placed under the control of appropriate regulatory sequences. The boundaries of the coding sequence are determined by a start codon at the 5′ (amino) terminus and a translation stop codon at the 3′ (carboxyl) terminus. A coding sequence can include, but is not limited to, prokaryotic sequences, cDNA from eukaryotic mRNA, genomic DNA sequences from eukaryotic (e.g., mammalian) DNA, and even synthetic DNA sequences. If the coding sequence is intended for expression in a eukaryotic cell, a polyadenylation signal and transcription termination sequence will usually be located 3′ to the coding sequence.

An ''Expression vector'' is a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest operably linked to additional segments that provide for its

transcription. Such additional segments may include promoter and terminator sequences, and optionally one or more origins of replication, one or more selectable markers, an enhancer, a polyadenylation signal, and the like. Expression vectors are generally derived from plasmid or viral DNA, or may contain elements of both.

Transcriptional and translational control sequences are DNA regulatory sequences, such as promoters, enhancers, terminators, and the like, that provide for the expression of a coding sequence in a host cell. In eukaryotic cells, polyadenylation signals are control sequences.

A ''secretory signal sequence'' is a DNA sequence that encodes a polypeptide (a ''secretory peptide'' that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

The term ''promoter'' is used herein for its art-recognized meaning to denote a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes.

''Operably linked'', when referring to DNA segments, indicates that the segments are arranged so that they function in concert for their intended purposes, e.g. transcription initiates in the promoter and proceeds through the coding segment to the terminator.

A coding sequence is ''under the control'' of transcriptional and translational control sequences in a cell when RNA polymerase transcribes the coding sequence into mRNA, which is then trans-RNA spliced and translated into the protein encoded by the coding sequence.

''Isolated polypeptide'' is a polypeptide which is essentially free of other non-[enzyme] polypeptides, e.g., at least about 20% pure, preferably at least about 40% pure, more preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by SDS-PAGE.

''Heterologous'' DNA refers to DNA not naturally located in the cell, or in a chromosomal site of the cell. Preferably, the heterologous DNA includes a gene foreign to the cell.

A cell has been ''transfected'' by exogenous or heterologous DNA when such DNA has been introduced inside the cell. A cell has been ''transformed'' by exogenous or heterologous DNA when the transfected DNA effects a phenotypic change. Preferably, the transforming DNA should be integrated (covalently linked) into chromosomal DNA making up the genome of the cell.

A ''clone'' is a population of cells derived from a single cell or common ancestor by mitosis.

''Homologous recombination'' refers to the insertion of a foreign DNA sequence of a vector in a chromosome. Preferably, the vector targets a specific chromosomal site for homologous recombination. For specific homologous recombination, the vector will contain sufficiently long regions of homology to se-

quences of the chromosome to allow complementary binding and incorporation of the vector into the chromosome. Longer regions of homology, and greater degrees of sequence similarity, may increase the efficiency of homologous recombination.

Nucleic Acid Sequence

The techniques used to isolate or clone a nucleic acid sequence encoding a polypeptide are known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of the nucleic acid sequences of the present invention from such genomic DNA can be effected, e.g., by using the well known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, e.g., Innis et al., 1990, A Guide to Methods and Application, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nuceic acid sequence-based amplification (NASBA) may be used. The nucleic acid sequence may be cloned from a strain producing the polypeptide, or from another related organism and thus, for example, may be an allelic or species variant of the polypeptide encoding region of the nucleic acid sequence.

The term ''isolated" nucleic acid sequence as used herein refers to a nucleic acid sequence which is essentially free of other nucleic acid sequences, e.g., at least about 20% pure, preferably at least about 40% pure, more preferably about 60% pure, even more preferably about 80% pure, most preferably about 90% pure, and even most preferably about 95% pure, as determined by agarose gel electorphoresis. For example, an isolated nucleic acid sequence can be obtained by standard

cloning procedures used in genetic engineering to relocate the nucleic acid sequence from its natural location to a different site where it will be reproduced. The cloning procedures may involve excision and isolation of a desired nucleic acid fragment comprising the nucleic acid sequence encoding the polypeptide, insertion of the fragment into a vector molecule, and incorporation of the recombinant vector into a host cell where multiple copies or clones of the nucleic acid sequence will be replicated. The nucleic acid sequence may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof.

Nucleic Acid Construct

As used herein the term ''nucleic acid construct'' is intended to indicate any nucleic acid molecule of cDNA, genomic DNA, synthetic DNA or RNA origin. The term "construct" is intended to indicate a nucleic acid segment which may be single- or double-stranded, and which may be based on a complete or partial naturally occurring nucleotide sequence encoding a polypeptide of interest. The construct may optionally contain other nucleic acid segments.

The DNA of interest may suitably be of genomic or cDNA origin, for instance obtained by preparing a genomic or cDNA library and screening for DNA sequences coding for all or part of the polypeptide by hybridization using synthetic oligonucleotide probes in accordance with standard techniques (cf. Sambrook et al., supra).

The nucleic acid construct may also be prepared synthetically by established standard methods, e.g. the phosphoamidite method described by Beaucage and Caruthers, Tetrahedron Let-

ters 22 (1981), 1859 - 1869, or the method described by Matthes et al., EMBO Journal 3 (1984), 801 - 805. According to the phosphoamidite method, oligonucleotides are synthesized, e.g. in an automatic DNA synthesizer, purified, annealed, ligated and cloned in suitable vectors.

Furthermore, the nucleic acid construct may be of mixed synthetic and genomic, mixed synthetic and cDNA or mixed genomic and cDNA origin prepared by ligating fragments of synthetic, genomic or cDNA origin (as appropriate), the fragments corresponding to various parts of the entire nucleic acid construct, in accordance with standard techniques.

The nucleic acid construct may also be prepared by polymerase chain reaction using specific primers, for instance as described in US 4,683,202 or Saiki et al., Science 239 (1988), 487 - 491.

The term nucleic acid construct may be synonymous with the term expression cassette when the nucleic acid construct contains all the control sequences required for expression of a coding sequence of the present invention. The term ''coding sequence'' as defined herein is a sequence which is transcribed into mRNA and translated into a polypeptide of the present invention when placed under the control of the above mentioned control sequences. The boundaries of the coding sequence are generally determined by a translation start codon ATG at the 5'-terminus and a translation stop codon at the 3'-terminus. A coding sequence can include, but is not limited to, DNA, cDNA, and recombinant nucleic acid sequences.

The term ''control sequences'' is defined herein to include all components which are necessary or advantageous for expression of the coding sequence of the nucleic acid sequence. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, a polyadenylation sequence, a propeptide sequence, a promoter, a signal sequence, and a transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleic acid sequence encoding a polypeptide.

The control sequence may be an appropriate promoter sequence, a nucleic acid sequence which is recognized by a host cell for expression of the nucleic acid sequence. The promoter sequence contains transcription and translation control sequences which mediate the expression of the polypeptide. The promoter may be any nucleic acid sequence which shows transcriptional activity in the host cell of choice and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell.

The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleic acid sequence encoding the polypeptide. Any terminator which is functional in the host cell of choice may be used

in the present invention.

The control sequence may also be a polyadenylation sequence, a sequence which is operably linked to the 3' terminus of the nucleic acid sequence and which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention.

The control sequence may also be a signal peptide coding region, which codes for an amino acid sequence linked to the amino terminus of the polypeptide which can direct the expressed polypeptide into the cell's secretory pathway of the host cell. The 5' end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region which encodes the secreted polypeptide. Alternatively, the 5' end of the coding sequence may contain a signal peptide coding region which is foreign to that portion of the coding sequence which encodes the secreted polypeptide. A foreign signal peptide coding region may be required where the coding sequence does not normally contain a signal peptide coding region. Alternatively, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in order to obtain enhanced secretion relative to the natural signal peptide coding region normally associated with the coding sequence. The signal peptide coding region may be obtained from a glucoamylase or an amylase gene from an Aspergillus species, a lipase or proteinase gene from a Rhizomucor species, the gene for the alpha-factor from Saccharomyces cerevisiae, an amylase or a protease gene from a Bacillus species, or the calf preprochymosin gene.

However, any signal peptide coding region capable of directing the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention.

The control sequence may also be a propeptide coding region, which codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a proenzyme or propolypeptide (or a zymogen in some cases). A propolypeptide is generally inactive and can be converted to mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the propolypeptide. The propeptide coding region may be obtained from the Bacillus subtilis alkaline protease gene (aprE), the Bacillus subtilis neutral protease gene (nprT), the Saccharomyces cerevisiae alpha-factor gene, or the Myceliophthora thermophilum laccase gene (WO 95/33836).

The nucleic acid constructs of the present invention may also comprise one or more nucleic acid sequences which encode one or more factors that are advantageous in the expression of the polypeptide, e.g., an activator (e.g., a trans-acting factor), a chaperone, and a processing protease. Any factor that is functional in the host cell of choice may be used in the present invention. The nucleic acids encoding one or more of these factors are not necessarily in tandem with the nucleic acid sequence encoding the polypeptide.

An activator is a protein which activates transcription of a nucleic acid sequence encoding a polypeptide (Kudla et al., 1990, EMBO Journal 9:1355-1364; Jarai and Buxton, 1994, Current Genetics 26:2238-244; Verdier, 1990, Yeast 6:271-297). The nucleic acid sequence encoding an activator may be obtained from the genes encoding Bacillus stearothermophilus

NprA (nprA), Saccharomyces cerevisiae heme activator protein 1 (hap1), Saccharomyces cerevisiae galactose metabolizing protein 4 (gal4), and Aspergillus nidulans ammonia regulation protein (areA). For further examples, see Verdier, 1990, supra and MacKenzie et al., 1993, Journal of General Microbiology 139:2295-2307.

A chaperone is a protein which assists another polypeptide in folding properly (Hartl et al., 1994, TIBS 19:20-25; Bergeron et al., 1994, TIBS 19:124-128; Demolder et al., 1994, Journal of Biotechnology 32:179-189; Craig, 1993, Science 260:1902-1903; Gething and Sambrook, 1992, Nature 355:33-45; Puig and Gilbert, 1994, Journal of Biological Chemistry 269:7764-7771; Wang and Tsou, 1993, The FASEB Journal 7:1515-11157; Robinson et al., 1994, Bio/Technology 1:381-384). The nucleic acid sequence encoding a chaperone may be obtained from the genes encoding Bacillus subtilis GroE proteins, Aspergillus oryzae protein disulphide isomerase, Saccharomyces cerevisiae calnexin, Saccharomyces cerevisiae BiP/GRP78, and Saccharomyces cerevisiae Hsp70. For further examples, see Gething and Sambrook, 1992, supra, and Hartl et al., 1994, supra.

A processing protease is a protease that cleaves a propeptide to generate a mature biochemically active polypeptide (Enderlin and Ogrydziak, 1994, Yeast 10:67-79; Fuller et al., 1989, Proceedings of the National Academy of Sciences USA 86:1434-1438; Julius et al., 1984, Cell 37:1075-1089; Julius et al., 1983, Cell 32:839-852). The nucleic acid sequence encoding a processing protease may be obtained from the genes encoding Aspergillus niger Kex2, Saccharomyces cerevisiae dipeptidylaminopeptidase, Saccharomyces cerevisiae Kex2, and Yarrowia lipolytica dibasic processing endoprotease (xpr6).

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. Regulatory systems in prokaryotic systems would include the lac, tac, and trp operator systems. In yeast, the ADH2 system or GAL1 system may be used. In filamentous fungi, the TAKA alpha-amylase promoter, Aspergillus niger glucoamylase promoter, and the Aspergillus oryzae glucoamylase promoter may be used as regulatory sequences. Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene which is amplified in the presence of methotrexate, and the metallothionein genes which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the polypeptide would be placed in tandem with the regulatory sequence.

Promoters

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention, especially in a bacterial host cell, are the promoters obtained from the E. coli lac operon, the Streptomyces coelicolor agarase gene (dagA), the Bacillus subtilis levansucrase gene (sacB), the Bacillus subtilis alkaline protease gene, the Bacillus licheniformis alpha-amylase gene (amyL), the Bacillus stearothermophilus maltogenic amylase gene (amyM), the Bacillus amyloliquefaciens alpha-amylase gene (amyQ), the Bacillus amyloliquefaciens BAN amylase gene, the Bacillus licheniformis

penicillinase gene (penP), the Bacillus subtilis xylA and xylB genes, and the prokaryotic beta-lactamase gene (Villa-Kamaroff et al., 1978, Proceedings of the National Academy of Sciences USA 75:3727-3731), as well as the tac promoter (DeBoer et al., 1983, Proceedings of the National Academy of Sciences USA 80:21-25) , or the Bacillus pumilus xylosidase gene, or by the phage Lambda PR or PL promoters or the E. coli lac, trp or tac promoters. Further promoters are described in "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242:74-94; and in Sambrook et al., 1989, supra.

Examples of suitable promoters for directing the transcription of the nucleic acid constructs of the present invention in a filamentous fungal host cell are promoters obtained from the genes encoding Aspergillus oryzae TAKA amylase, Rhizomucor miehei aspartic proteinase, Aspergillus niger neutral alpha-amylase, Aspergillus niger acid stable alpha-amylase, Aspergillus niger or Aspergillus awamori glucoamylase (glaA), Rhizomucor miehei lipase, Aspergillus oryzae alkaline protease, Aspergillus oryzae triose phosphate isomerase, Aspergillus nidulans acetamidase, Fusarium oxysporum trypsin-like protease (as described in U.S. Patent No. 4,288,627, which is incorporated herein by reference), and hybrids thereof. Particularly preferred promoters for use in filamentous fungal host cells are the TAKA amylase, NA2-tpi (a hybrid of the promoters from the genes encoding Aspergillus niger neutral (-amylase and Aspergillus oryzae triose phosphate isomerase), and glaA promoters. Further suitable promoters for use in filamentous fungus host cells are the ADH3 promoter (McKnight et al., The EMBO J. 4 (1985), 2093 - 2099) or the tpiA promoter.

Examples of suitable promoters for use in yeast host cells include promoters from yeast glycolytic genes (Hitzeman et al., J. Biol. Chem. 255 (1980), 12073 - 12080; Alber and Kawasaki, J. Mol. Appl. Gen. 1 (1982), 419 - 434) or alcohol dehydrogenase genes (Young et al., in Genetic Engineering of Microorganisms for Chemicals (Hollaender et al, eds.), Plenum Press, New York, 1982), or the TPI1 (US 4,599,311) or ADH2-4c (Russell et al., Nature 304 (1983), 652 - 654) promoters.

Further useful promoters are obtained from the Saccharomyces cerevisiae enolase (ENO-1) gene, the Saccharomyces cerevisiae galactokinase gene (GAL1), the Saccharomyces cerevisiae alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase genes (ADH2/GAP), and the Saccharomyces cerevisiae 3-phosphoglycerate kinase gene. Other useful promoters for yeast host cells are described by Romanos et al., 1992, Yeast 8:423-488. In a mammalian host cell, useful promoters include viral promoters such as those from Simian Virus 40 (SV40), Rous sarcoma virus (RSV), adenovirus, and bovine papilloma virus (BPV).

Examples of suitable promoters for directing the transcription of the DNA encoding the polypeptide of the invention in mammalian cells are the SV40 promoter (Subramani et al., Mol. Cell Biol. 1 (1981), 854 -864), the MT-1 (metallothionein gene) promoter (Palmiter et al., Science 222 (1983), 809 - 814) or the adenovirus 2 major late promoter.
An example of a suitable promoter for use in insect cells is the polyhedrin promoter (US 4,745,051; Vasuvedan et al., FEBS Lett. 311, (1992) 7 - 11), the P10 promoter (J.M. Vlak et al., J. Gen. Virology 69, 1988, pp. 765-776), the Autographa californica polyhedrosis virus basic protein promoter (EP 397

485), the baculovirus immediate early gene 1 promoter (US 5,155,037; US 5,162,222), or the baculovirus 39K delayed-early gene promoter (US 5,155,037; US 5,162,222).

Terminators

Preferred terminators for filamentous fungal host cells are obtained from the genes encoding Aspergillus oryzae TAKA amylase, Aspergillus niger glucoamylase, Aspergillus nidulans anthranilate synthase, Aspergillus niger alpha-glucosidase, and Fusarium oxysporum trypsin-like protease. for fungal hosts) the TPI1 (Alber and Kawasaki, op. cit.) or ADH3 (McKnight et al., op. cit.) terminators.

Preferred terminators for yeast host cells are obtained from the genes encoding Saccharomyces cerevisiae enolase, Saccharomyces cerevisiae cytochrome C (CYC1), or Saccharomyces cerevisiae glyceraldehyde-3-phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos et al., 1992, supra.

Polyadenylation Signals

Preferred polyadenylation sequences for filamentous fungal host cells are obtained from the genes encoding Aspergillus oryzae TAKA amylase, Aspergillus niger glucoamylase, Aspergillus nidulans anthranilate synthase, and Aspergillus niger alpha-glucosidase.

Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, Molecular Cellular Biology 15:5983-5990.

Polyadenylation sequences are well known in the art for mammalian host cells such as SV40 or the adenovirus 5 Elb region.

Signal Sequences

An effective signal peptide coding region for bacterial host cells is the signal peptide coding region obtained from the maltogenic amylase gene from Bacillus NCIB 11837, the Bacillus stearothermophilus alpha-amylase gene, the Bacillus licheniformis subtilisin gene, the Bacillus licheniformis beta-lactamase gene, the Bacillus stearothermophilus neutral proteases genes (nprT, nprS, nprM), and the Bacillus subtilis PrsA gene. Further signal peptides are described by Simonen and Palva, 1993, Microbiological Reviews 57:109-137.

An effective signal peptide coding region for filamentous fungal host cells is the signal peptide coding region obtained from Aspergillus oryzae TAKA amylase gene, Aspergillus niger neutral amylase gene, the Rhizomucor miehei aspartic proteinase gene, the Humicola lanuginosa cellulase or lipase gene, or the Rhizomucor miehei lipase or protease gene, Aspergillus sp. amylase or glucoamylase, a gene encoding a Rhizomucor miehei lipase or protease. The signal peptide is preferably derived from a gene encoding A. oryzae TAKA amylase, A. niger neutral (-amylase, A. niger acid-stable amylase, or A. niger glucoamylase.

Useful signal peptides for yeast host cells are obtained from the genes for Saccharomyces cerevisiae a-factor and Saccharomyces cerevisiae invertase. Other useful signal peptide coding regions are described by Romanos et al., 1992, supra.
For secretion from yeast cells, the secretory signal sequence may encode any signal peptide which ensures efficient direction of the expressed polypeptide into the secretory pathway of the cell. The signal peptide may be naturally occurring signal peptide, or a functional part thereof, or it may be a synthetic peptide. Suitable signal peptides have been found to

be the a-factor signal peptide (cf. US 4,870,008), the signal peptide of mouse salivary amylase (cf. O. Hagenbuchle et al., Nature 289, 1981, pp. 643-646), a modified carboxypeptidase signal peptide (cf. L.A. Valls et al., Cell 48, 1987, pp. 887-897), the yeast BAR1 signal peptide (cf. WO 87/02670), or the yeast aspartic protease 3 (YAP3) signal peptide (cf. M. Egel-Mitani et al., Yeast 6, 1990, pp. 127-137).

For efficient secretion in yeast, a sequence encoding a leader peptide may also be inserted downstream of the signal sequence and uptream of the DNA sequence encoding the polypeptide. The function of the leader peptide is to allow the expressed polypeptide to be directed from the endoplasmic reticulum to the Golgi apparatus and further to a secretory vesicle for secretion into the culture medium (i.e. exportation of the polypeptide across the cell wall or at least through the cellular membrane into the periplasmic space of the yeast cell). The leader peptide may be the yeast a-factor leader (the use of which is described in e.g. US 4,546,082, EP 16 201, EP 123 294, EP 123 544 and EP 163 529). Alternatively, the leader peptide may be a synthetic leader peptide, which is to say a leader peptide not found in nature. Synthetic leader peptides may, for instance, be constructed as described in WO 89/02463 or WO 92/11378.

For use in insect cells, the signal peptide may conveniently be derived from an insect gene (cf. WO 90/05783), such as the lepidopteran Manduca sexta adipokinetic hormone precursor signal peptide (cf. US 5,023,328).

Expression Vectors

The present invention also relates to recombinant expression vectors comprising a nucleic acid sequence of the present in-

vention, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present invention may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression, and possibly secretion.

The recombinant expression vector may be any vector (e.g., a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleic acid sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids. The vector may be an autonomously replicating vector, i.e., a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. The vector system may be a single vector or plasmid or two or more vectors or plasmids which

together contain the total DNA to be introduced into the genome of the host cell, or a transposon.

The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like. Examples of bacterial selectable markers are the dal genes from Bacillus subtilis or Bacillus licheniformis, or markers which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol, tetracycline, neomycin, hygromycin or methotrexate resistance. A frequently used mammalian marker is the dihydrofolate reductase gene (DHFR). Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. A selectable marker for use in a filamentous fungal host cell may be selected from the group including, but not limited to, amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hygB (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase), sC (sulfate adenyltransferase), trpC (anthranilate synthase), and glufosinate resistance markers, as well as equivalents from other species. Preferred for use in an Aspergillus cell are the amdS and pyrG markers of Aspergillus nidulans or Aspergillus oryzae and the bar marker of Streptomyces hygroscopicus. Furthermore, selection may be accomplished by co-transformation, e.g., as described in WO 91/17243, where the selectable marker is on a separate vector.

The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector into

the host cell genome or autonomous replication of the vector in the cell independent of the genome of the cell.

The vectors of the present invention may be integrated into the host cell genome when introduced into a host cell. For integration, the vector may rely on the nucleic acid sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination. These nucleic acid sequences may be any sequence that is homologous with a target sequence in the genome of the host cell, and, furthermore, may be non-encoding or encoding sequences.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autono-

mously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322, pUC19, pACYC177, pACYC184, pUB110, pE194, pTA1060, and pAMß1. Examples of origin of replications for use in a yeast host cell are the 2 micron origin of replication, the combination of CEN6 and ARS4, and the combination of CEN3 and ARS1. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, e.g., Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75:1433).

More than one copy of a nucleic acid sequence encoding a polypeptide of the present invention may be inserted into the host cell to amplify expression of the nucleic acid sequence. Stable amplification of the nucleic acid sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome using methods well known in the art and selecting for transformants.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, e.g., Sambrook et al., 1989, supra).

Host Cells

The present invention also relates to recombinant host cells, comprising a nucleic acid sequence of the invention, which are advantageously used in the recombinant production of the polypeptides. The term "host cell" encompasses any progeny of a parent cell which is not identical to the parent cell due to mutations that occur during replication.

The cell is preferably transformed with a vector comprising a nucleic acid sequence of the invention followed by integration of the vector into the host chromosome. "Transformation" means introducing a vector comprising a nucleic acid sequence of the present invention into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector. Integration is generally considered to be an advantage as the nucleic acid sequence is more likely to be stably maintained in the cell. Integration of the vector into the host chromosome may occur by homologous or non-homologous recombination as described above.

The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source. The host cell may be a unicellular microorganism, e.g., a prokaryote, or a non-unicellular microorganism, e.g., a eukaryote. Useful unicellular cells are bacterial cells such as gram positive bacteria including, but not limited to, a Bacillus cell, e.g., Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus brevis, Bacillus circulans, Bacillus coagulans, Bacillus lautus, Bacillus lentus, Bacillus licheniformis, Bacillus megaterium, Bacillus stearothermophilus, Bacillus subtilis, and Bacillus thuringiensis; or a Streptomyces cell, e.g., Streptomyces lividans or Streptomyces murinus, or gram negative bacteria such as E. coli and Pseudomonas sp. In a preferred embodiment, the bacterial host cell is a Bacillus lentus, Bacillus licheniformis, Bacillus stearothermophilus or Bacillus subtilis cell. The transformation of a bacterial host cell may, for instance, be effected by protoplast transformation (see, e.g., Chang and Cohen, 1979, Molecular General Genetics 168:111-115), by using competent cells (see, e.g.,

Young and Spizizin, 1961, Journal of Bacteriology 81:823-829, or Dubnar and Davidoff-Abelson, 1971, Journal of Molecular Biology 56:209-221), by electroporation (see, e.g., Shigekawa and Dower, 1988, Biotechniques 6:742-751), or by conjugation (see, e.g., Koehler and Thorne, 1987, Journal of Bacteriology 169:5771-5278).

The host cell may be a eukaryote, such as a mammalian cell, an insect cell, a plant cell or a fungal cell.
Useful mammalian cells include Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, COS cells, or any number of other immortalized cell lines available, e.g., from the American Type Culture Collection.

Examples of suitable mammalian cell lines are the COS (ATCC CRL 1650 and 1651), BHK (ATCC CRL 1632, 10314 and 1573, ATCC CCL 10), CHL (ATCC CCL39) or CHO (ATCC CCL 61) cell lines. Methods of transfecting mammalian cells and expressing DNA sequences introduced in the cells are described in e.g. Kaufman and Sharp, J. Mol. Biol. 159 (1982), 601 - 621; Southern and Berg, J. Mol. Appl. Genet. 1 (1982), 327 - 341; Loyter et al., Proc. Natl. Acad. Sci. USA 79 (1982), 422 - 426; Wigler et al., Cell 14 (1978), 725; Corsaro and Pearson, Somatic Cell Genetics 7 (1981), 603, Ausubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, Inc., N.Y., 1987, Hawley-Nelson et al., Focus 15 (1993), 73; Ciccarone et al., Focus 15 (1993), 80; Graham and van der Eb, Virology 52 (1973), 456; and Neumann et al., EMBO J. 1 (1982), 841 - 845.

In a preferred embodiment, the host cell is a fungal cell. "Fungi" as used herein includes the phyla Ascomycota, Basidiomycota, Chytridiomycota, and Zygomycota (as defined by

Hawksworth et al., In, Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) as well as the Oomycota (as cited in Hawksworth et al., 1995, supra, page 171) and all mitosporic fungi (Hawksworth et al., 1995, supra). Representative groups of Ascomycota include, e.g., Neurospora, Eupenicillium (=Penicillium), Emericella (=Aspergillus), Eurotium (=Aspergillus), and the true yeasts listed above. Examples of Basidiomycota include mushrooms, rusts, and smuts. Representative groups of Chytridiomycota include, e.g., Allomyces, Blastocladiella, Coelomomyces, and aquatic fungi. Representative groups of Oomycota include, e.g., Saprolegniomycetous aquatic fungi (water molds) such as Achlya. Examples of mitosporic fungi include Aspergillus, Penicillium, Candida, and Alternaria. Representative groups of Zygomycota include, e.g., Rhizopus and Mucor.

In a preferred embodiment, the fungal host cell is a yeast cell. "Yeast" as used herein includes ascosporogenous yeast (Endomycetales), basidiosporogenous yeast, and yeast belonging to the Fungi Imperfecti (Blastomycetes). The ascosporogenous yeasts are divided into the families Spermophthoraceae and Saccharomycetaceae. The latter is comprised of four subfamilies, Schizosaccharomycoideae (e.g., genus Schizosaccharomyces), Nadsonioideae, Lipomycoideae, and Saccharomycoideae (e.g., genera Pichia, Kluyveromyces and Saccharomyces). The basidiosporogenous yeasts include the genera Leucosporidim, Rhodosporidium, Sporidiobolus, Filobasidium, and Filobasidiella. Yeast belonging to the Fungi Imperfecti are divided into two families, Sporobolomycetaceae (e.g., genera Sorobolomyces and Bullera) and Cryptococcaceae (e.g., genus Candida). Since the classification of yeast may change in the future, for the purposes of this invention, yeast shall be defined as described in Biology and Activities of Yeast (Skinner, F.A.,

Passmore, S.M., and Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series No. 9, 1980. The biology of yeast and manipulation of yeast genetics are well known in the art (see, e.g., Biochemistry and Genetics of Yeast, Bacil, M., Horecker, B.J., and Stopani, A.O.M., editors, 2nd edition, 1987; The Yeasts, Rose, A.H., and Harrison, J.S., editors, 2nd edition, 1987; and The Molecular Biology of the Yeast Saccharomyces, Strathern et al., editors, 1981).

The yeast host cell may be selected from a cell of a species of Candida, Kluyveromyces, Saccharomyces, Schizosaccharomyces, Candida, Pichia, Hansehula, , or Yarrowia. In a preferred embodiment, the yeast host cell is a Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis or Saccharomyces oviformis cell. Other useful yeast host cells are a Kluyveromyces lactis Kluyveromyces fragilis Hansehula polymorpha, Pichia pastoris Yarrowia lipolytica, Schizosaccharomyces pombe, Ustilgo maylis, Candida maltose, Pichia guillermondii and Pichia methanolio cell (cf. Gleeson et al., J. Gen. Microbiol. 132, 1986, pp. 3459-3465; US 4,882,279 and US 4,879,231).

In a preferred embodiment, the fungal host cell is a filamentous fungal cell. "Filamentous fungi" include all filamentous forms of the subdivision Eumycota and Oomycota (as defined by Hawksworth et al., 1995, supra). The filamentous fungi are characterized by a vegetative mycelium composed of chitin, cellulose, glucan, chitosan, mannan, and other complex polysaccharides. Vegetative growth is by hyphal elongation and carbon catabolism is obligately aerobic. In contrast, vegetative growth by yeasts such as Saccharomyces cerevisiae

is by budding of a unicellular thallus and carbon catabolism may be fermentative. In a more preferred embodiment, the filamentous fungal host cell is a cell of a species of, but not limited to, Acremonium, Aspergillus, Fusarium, Humicola, Mucor, Myceliophthora, Neurospora, Penicillium, Thielavia, Tolypocladium, and Trichoderma or a teleomorph or synonym thereof. In an even more preferred embodiment, the filamentous fungal host cell is an Aspergillus cell. In another even more preferred embodiment, the filamentous fungal host cell is an Acremonium cell. In another even more preferred embodiment, the filamentous fungal host cell is a Fusarium cell. In another even more preferred embodiment, the filamentous fungal host cell is a Humicola cell. In another even more preferred embodiment, the filamentous fungal host cell is a Mucor cell. In another even more preferred embodiment, the filamentous fungal host cell is a Myceliophthora cell. In another even more preferred embodiment, the filamentous fungal host cell is a Neurospora cell. In another even more preferred embodiment, the filamentous fungal host cell is a Penicillium cell. In another even more preferred embodiment, the filamentous fungal host cell is a Thielavia cell. In another even more preferred embodiment, the filamentous fungal host cell is a Tolypocladium cell. In another even more preferred embodiment, the filamentous fungal host cell is a Trichoderma cell. In a most preferred embodiment, the filamentous fungal host cell is an Aspergillus awamori, Aspergillus foetidus, Aspergillus japonicus, Aspergillus niger, Aspergillus nidulans or Aspergillus oryzae cell. In another most preferred embodiment, the filamentous fungal host cell is a Fusarium cell of the section Discolor (also known as the section Fusarium). For example, the filamentous fungal parent cell may be a Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium

culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochroum, Fusarium sulphureum, or Fusarium trichothecioides cell. In another prefered embodiment, the filamentous fungal parent cell is a Fusarium strain of the section Elegans, e.g., Fusarium oxysporum. In another most preferred embodiment, the filamentous fungal host cell is a Humicola insolens or Humicola lanuginosa cell. In another most preferred embodiment, the filamentous fungal host cell is a Mucor miehei cell. In another most preferred embodiment, the filamentous fungal host cell is a Myceliophthora thermophilum cell. In another most preferred embodiment, the filamentous fungal host cell is a Neurospora crassa cell. In another most preferred embodiment, the filamentous fungal host cell is a Penicillium purpurogenum cell. In another most preferred embodiment, the filamentous fungal host cell is a Thielavia terrestris cell or a Acremonium chrysogenum cell. In another most preferred embodiment, the Trichoderma cell is a Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei or Trichoderma viride cell. The use of Aspergillus spp. for the expression of proteins is described in, e.g., EP 272 277, EP 230 023.

Transformation

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known per se. Suitable procedures for transformation of Aspergillus host cells are described in EP 238 023 and Yelton et al., 1984, Proceedings of the National Academy of Sciences USA 81:1470-1474. A suitable method of transforming Fusarium species is described by

Malardier et al., 1989, Gene 78:147-156 or in copending US Serial No. 08/269,449. Examples of other fungal cells are cells of filamentous fungi, e.g. Aspergillus spp., Neurospora spp., Fusarium spp. or Trichoderma spp., in particular strains of A. oryzae, A. nidulans or A. niger. The use of Aspergillus spp. for the expression of proteins is described in, e.g., EP 272 277, EP 230 023, EP 184 ... The transformation of F. oxysporum may, for instance, be carried out as described by Malardier et al., 1989, Gene 78: 147-156.

Yeast may be transformed using the procedures described by Becker and Guarente, In Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., 1983, Journal of Bacteriology 153:163; and Hinnen et al., 1978, Proceedings of the National Academy of Sciences USA 75:1920. Mammalian cells may be transformed by direct uptake using the calcium phosphate precipitation method of Graham and Van der Eb (1978, Virology 52:546). Transformation of insect cells and production of heterologous polypeptides therein may be performed as described in US 4,745,051; US 4, 775, 624; US 4,879,236; US 5,155,037; US 5,162,222; EP 397,485) all of which are incorporated herein by reference. The insect cell line used as the host may suitably be a Lepidoptera cell line, such as Spodoptera frugiperda cells or Trichoplusia ni cells (cf. US 5,077,214). Culture conditions may suitably be as described in, for instance, WO 89/01029 or WO 89/01028, or any of the aforementioned references.

Methods of Production

The transformed or transfected host cells described above are cultured in a suitable nutrient medium under conditions permitting the production of the desired molecules, after which these are recovered from the cells, or the culture broth.

The medium used to culture the cells may be any conventional medium suitable for growing the host cells, such as minimal or complex media containing appropriate supplements. Suitable media are available from commercial suppliers or may be prepared according to published recipes (e.g. in catalogues of the American Type Culture Collection). The media are prepared using procedures known in the art (see, e.g., references for bacteria and yeast; Bennett, J.W. and LaSure, L., editors, More Gene Manipulations in Fungi, Academic Press, CA, 1991).

If the molecules are secreted into the nutrient medium, they can be recovered directly from the medium. If they are not secreted, they can be recovered from cell lysates. The molecules are recovered from the culture medium by conventional procedures including separating the host cells from the medium by centrifugation or filtration, precipitating the proteinaceous components of the supernatant or filtrate by means of a salt, e.g. ammonium sulphate, purification by a variety of chromatographic procedures, e.g. ion exchange chromatography, gelfiltration chromatography, affinity chromatography, or the like, dependent on the type of molecule in question.

The molecules of interest may be detected using methods known in the art that are specific for the molecules. These detection methods may include use of specific antibodies, formation of a product, or disappearance of a substrate. For example, an enzyme assay may be used to determine the activity of the

molecule. Procedures for determining various kinds of activity are known in the art.

The molecules of the present invention may be purified by a variety of procedures known in the art including, but not limited to, chromatography (e.g., ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (e.g., preparative isoelectric focusing (IEF), differential solubility (e.g., ammonium sulfate precipitation), or extraction (see, e.g., Protein Purification, J-C Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

The term ''immunological response'', used in connection with the present invention, is the response of an organism to a compound, which involves the immune system according to any of the four standard reactions (Type I, II, III and IV according to Coombs & Gell).

Correspondingly, the ''immunogenicity'' of a compound used in connection with the present invention refers to the ability of this compound to induce an 'immunological response' in animals including man.

The term ''allergic response'', used in connection with the present invention, is the response of an organism to a compound, which involves IgE mediated responses (Type I reaction according to Coombs & Gell). It is to be understood that sensibilization (i.e. development of compound-specific IgE antibodies) upon exposure to the compound is included in the definition of ''allergic response''.

Correspondingly, the ''allergenicity'' of a compound used in connection with the present invention refers to the ability of this compound to induce an 'allergic response' in animals including man.

The term "parent protein'' refer to the polypeptide to be modified by creating a library of diversified mutants. The ''parent protein'' may be a naturally occurring (or wild-type) polypeptide or it may be a variant thereof prepared by any suitable means. For instance, the ''parent protein'' may be a variant of a naturally occurring polypeptide which has been modified by substitution, deletion or truncation of one or more amino acid residues or by addition or insertion of one or more amino acid residues to the amino acid sequence of a naturally-occurring polypeptide.

The term '' randomized library'' of protein variants refers to a library with at least partially randomized composition of the members, e.g. protein variants.

An ''epitope'' is a set of amino acids on a protein that are involved in an immunological response, such as antibody binding or T-cell activation. One particularly useful method of identifying epitopes involved in antibody binding is to screen a library of peptide-phage membrane protein fusions and selecting those that bind to relevant antigen-specific antibodies, sequencing the randomized part of the fusion gene, aligning the sequences involved in binding, defining consensus sequences based on these alignments, and mapping these consensus sequences on the surface or the sequence and/or structure of the antigen, to identify epitopes involved in antibody binding.

By the term ''epitope pattern'' is meant such a consensus sequence of antibody binding peptides. An example is the epitope pattern A R R < R. The sign ''<'' in this notation indicates that the aligned antibody binding peptides included a non-consensus amino acid between the second and the third arginine.

An ''epitope area'' is defined as the amino acids situated close to the epitope sequence amino acids. Preferably, the amino acids of an epitope area are located <5Å from the epitope sequence. Hence, an epitope area also includes the corresponding epitope sequence itself. Modifications of amino acids of the 'epitope area' can possibly affect the immunogenic function of the corresponding epitope.

By the term ''epitope sequence'' is meant the amino acid residues of a parent protein, which have been identified to belong to an epitope by the methods of the present invention (an example of an epitope sequence is E271 Q12 I8 in Savinase).

The term 'antibody binding peptide' denotes a peptide that bind with sufficiently high affinity to antibodies. Identification of 'antibody binding peptides' and their sequences constitute the first step of the method of this invention.

''Anchor amino acids'' are the individual amino acids of an epitope pattern.

''Hot spot amino acids'' are amino acids of parent protein, which are particularly likely to result in modified immunogenecity if they are mutated. Amino acids, which appear in three

or more epitope sequences or which correspond to anchor amino acids are hot spot amino acids.

''Environmental allergens'' are protein allergens that are present naturally. They include pollen, dust mite allergens, pet allergens, food allergens, venoms, etc.

''Commercial allergens'' are protein allergens that are being brought to the market commercially. They include enzymes, pharmaceutical proteins, antimicrobial peptides, as well as allergens of transgenic plants.

The ''donor protein'' is the protein that was used to raise antibodies used to identify antibody binding sequences, hence the donor protein provides the information that leads to the epitope patterns.

The ''acceptor protein'' is the protein, whose structure is used to fit the identified epitope patterns and/or to fit the antibody binding sequences. Hence the acceptor protein is also the parent protein.

An ''autoepitope'' is one that has been identified using antibodies raised against the parent protein, i.e. the acceptor and the donor proteins are identical.

A ''heteroepitope'' is one that has been identified with distinct donor and acceptor proteins.

The term ''functionality'' of protein variants refers to e.g. enzymatic activity; binding to a ligand or receptor; stimula-

tion of a cellular response (e.g. ^{3}H-thymidine incorporation as response to a mitogenic factor); or anti-microbial activity.

By the term ''specific polyclonal antibodies'' is meant polyclonal antibodies isolated according to their specificity for a certain antigen, e.g. the protein backbone.

By the term ''monospecific antibodies'' is meant polyclonal antibodies isolated according to their specificity for a certain epitope. Such monospecific antibodies will bind to the same epitope, but with different affinity, as they are produced by a number of antibody producing cells recognizing overlapping but not necessarily identical epitopes.

The term ''randomized library'' of protein variants refers to a library with at least partially randomized composition of the members, e.g. protein variants.

'Spiked mutagenesis' is a form of site-directed mutagenesis, in which the primers used have been synthesized using mixtures of oligonucleotides at one or more positions.

By the term ''a protein variant having modified immunogenicity as compared to the parent protein'' is meant a protein variant which differs from the parent protein in one or more amino acids whereby the immunogenicity of the variant is modified. The modification of immunogenicity may be confirmed by testing the ability of the protein variant to elicit an IgE/IgG response.

In the present context the term ''protein'' is intended to cover oligopeptides, polypeptides as well as proteins as such.

Detailed description of the invention

The present invention relates to a method of selecting a protein variant having modified immunogenicity as compared to a parent protein,

comprising the steps of:

a) obtaining antibody binding peptide sequences,

b) using the sequences to localise epitope sequences on the 3-dimensional structure of parent protein,

c) defining an epitope area including amino acids situated within 5 Å from the epitope amino acids constituting the epitope sequence,

d) changing one or more of the amino acids defining the epitope area of the parent protein by genetic engineering mutations of a DNA sequence encoding the parent protein,

e) introducing the mutated DNA sequence into a suitable host, culturing said host and expressing the protein variant, and

f) evaluating the immunogenicity of the protein variant using the parent protein as reference.

A) How to find antibody binding peptide sequences and epitope patterns

A first step of the method is to identify peptide sequences, which bind specifically to antibodies.

Antibody binding peptide sequences can be found by testing a set of known peptide sequences for binding to antibodies raised against the donor protein. These sequences are typically selected, such that each represents a segment of the donor protein sequence (Mol. Immunol., 1992, vol. 29, pp.1383-1389; Am. J. Resp. Cell. Mol. Biol. 2000, vol. 22, pp. 344-351). Also, randomized synthetic peptide libraries can be used to find antibody binding sequences (Slootstra et al; Molecular Diversity, 1996, vol. 2, pp. 156-164).

In a preferred method, the identification of antibody binding sequences may be achieved by screening of a display package library, preferably a phage display library. The principle behind phage display is that a heterologous DNA sequence can be inserted in the gene coding for a coat protein of the phage (WO 92/15679). The phage will make and display the hybrid protein on its surface where it can interact with specific target agents. Such target agent may be antigen-specific antibodies. It is therefore possible to select specific phages that display antibody-binding peptide sequences. The displayed peptides can be of predetermined lengths, for example 9 amino acids long, with randomized sequences, resulting in a random peptide display package library. Thus, by screening for antibody binding, one can isolate the peptide sequences that have sufficiently high affinity for the particular antibody used. The peptides of the hybrid proteins of the specific phages

which bind protein-specific antibodies characterize epitopes that are recognized by the immune system.

The antibodies used for reacting with the display package are preferably IgE antibodies to ensure that the epitopes identified are IgE epitopes, i.e. epitopes inducing and binding IgE. In a preferred embodiment the antibodies are polyclonal antibodies, optionally monospecific antibodies.

For the purpose of the present invention polyclonal antibodies are preferred in order to obtain a broader knowledge about the epitopes of a protein.

It is of great importance that the amino acid sequence of the peptides presented by the display packages is long enough to represent a significant part of the epitope to be identified. In a preferred embodiment of the invention the peptides of the peptide display package library are oligopeptides having from 5 to 25 amino acids, preferably at least 8 amino acids, such as 9 amino acids. For a given length of peptide sequences (n), the theoretical number of different possible sequences can be calculated as 20^n. The diversity of the package library used must be large enough to provide a suitable representation of the theoretical number of different sequences. In a phage-display library, each phage has one specific sequence of a determined length. Hence an average phage display library can express 10^8 - 10^{12} different random sequences, and is therefore well-suited to represent the theoretical number of different sequences.

The antibody binding peptide sequences can be further analysed by consensus alignment e.g. by the methods described by Feng

and Doolittle, Meth. Enzymol., 1996, vol. 266, pp. 368-382; Feng and Doolittle, J. Mol. Evol., 1987, vol. 25, pp. 351-360; and Taylor,. Meth. Enzymol., 1996, vol. 266, pp. 343-367.

This leads to identification of epitope patterns, which can assist the comparison of the linear information obtained from the antibody binding peptide sequences to the 3-dimensional structure of the acceptor protein in order to identify epitope sequences at the surface of the acceptor protein.

B) How to identify epitope sequences and epitope areas.

Given a number of antibody binding peptide sequences and possibly the corresponding epitope patterns, one need the 3-dimensional structure coordinates of an acceptor protein to find the epitope sequences on its surface.

These coordinates can be found in databases (NCBI: http://www.ncbi.nlm.nih.gov/), determined experimentally using conventional methods (Ducruix and Giegé: Crystallization of Nucleic Acids and Proteins, IRL PRess, Oxford, 1992, ISBN 0-19-963245-6), or they can be deduced from the coordinates of a homologous protein. Typical actions required for the construction of a model structure are: alignment of homologous sequences for which 3-dimensional structures exist, definition of Structurally Conserved Regions (SCRs), assignment of coordinates to SCRs, search for structural fragments/loops in structure databases to replace Variable Regions, assignment of coordinates to these regions, and structural refinement by energy minimization. Regions containing large inserts (>3 resi-

dues) relative to the known 3-dimensional structures are known to be quite difficult to model, and structural predictions must be considered with care.

Using the coordinates and the several methods of mapping the linear information on the 3-dimensional surface are possible, as described in the examples below.

One can match each amino acid residue of the antibody binding peptide to an identical or homologous amino acid on the 3-D surface of the acceptor protein, such that amino acids that are adjacent in the primary sequence are close on the surface of the acceptor protein, with close being <5Å, preferably <3Å between any two atoms of the two amino acids.

Alternatively, one can define a geometric body (e.g. an ellipsoid, a sphere, or a box) of a size that matches a possible binding interface between antibody and antigen and look for a positioning of this body where it will contain most of or all the anchor amino acids.

Also, one can use the epitope patterns to facilitate identification of epitope sequences. This can be done, by first matching the anchor amino acids on the 3-D structure and subsequently looking for other elements of the antibody binding peptide sequences, which provide additional matches. If there are many residues to be matched, it is only necessary that a suitable number can be found on the 3-D structure. For example if an epitope pattern comprises 4, 5, 6, or 7 amino acids, it is only necessary that 3 matches surface elements of the acceptor protein..

In all cases, it is desirable that amino acids of the epitope sequence are surface exposed (as described below in Examples).

It is known, that amino acids that surround binding sequences can affect binding of a ligand without participating actively in the binding process. Based on this knowledge, areas covered by amino acids with potential steric effects on the epitope-antibody interaction, were defined around the identified epitope sequences. These areas are called 'epitope areas'. Practically, all amino acids situated within 5Å from the amino acids defining the epitope sequence were included. Preferably, the epitope area equals the epitope sequence. The accessibility criterium was not used as hidden amino acids of an epitope area also can have an effect on the adjacent amino acids of the epitope sequence.

C) How to use the epitope information.

There are at least four ways to utilize the information about epitope sequences, which has been derived by the methods of this invention:

1) reduce the allergenicity of a commercial protein using protein engineering.
2) reduce the potential of commercial proteins to cross-react with environmental allergens and hence cause allergic reactions in people sensitized to the environmental allergens (or vice versa).
3) improve the immunotherapeutic effect of allergen vaccines.
4) assist characterization of clinical allergies in order to select the appropriate allergen vaccine.

Protein engineering to reduce the allergenicity, cross-reactivity and/or immunotherapeutic effect of proteins.

The methods described thus far have led to identification of epitope areas on an acceptor protein, each containing epitope sequences. These subsets of amino acids, are preferred for introducing mutations that are meant to modify the immunogenecity of the acceptor protein. An even more preferred subset of amino acids to target by mutagenesis are 'hot spot amino acids', which appear in several different epitope sequences, or which corresponds to anchor amino acids of the epitope patterns.
Thus, genetic engineering mutations should be designed in the epitope areas, preferably in epitope sequences, and more preferably in the 'hot spot amino acids'.

Substitution, deletion, insertion

When the epitope area(s) have been identified, a protein variant exhibiting a modified immunogenicity may be produced by changing the identified epitope area of the parent protein by genetic engineering mutation of a DNA sequence encoding the parent protein.

The epitope identified may be changed by substituting at least one amino acid of the epitope area. In a preferred embodiment at least one anchor amino acid or hot spot amino acid is changed. The change will often be substituting to an amino acid of different size, hydrophilicity, and/or polarity, such

as a small amino acid versus a large amino acid, a hydrophilic amino acid versus a hydrophobic amino acid, a polar amino acid versus a non-polar amino acid and a basic versus an acidic amino acid.

Other changes may be the addition or deletion of at least one amino acid of the epitope sequence, preferably deleting an anchor amino acid or a hot spot amino acid. Furthermore, an epitope pattern may be changed by substituting some amino acids, and deleting/adding other.

When one uses protein engineering to eliminate epitopes, it is indeed possible that new epitopes are created, or existing epitopes are duplicated. To reduce this risk, one can map the planned mutations at a given position on the 3-dimensional structure of the protein of interest, and control the emerging amino acid constellation against a database of known epitope patterns, to rule out those possible replacement amino acids, which are predicted to result in creation or duplication of epitopes. Thus, risk mutations can be identified and eliminated by this procedure, thereby reducing the risk of making mutations that lead to increased rather than decreased allergenicity.

Introduction of residues for chemical derivatization in epitope areas

In yet another embodiment, one can design the mutation, such that amino acids suitable for chemical modification are substituted for existing ones in the epitope areas. The protein variant can then be conjugated to activated polymers. Which

amino acids to substitute and/or insert, depends in principle on the coupling chemistry to be applied. The chemistry for preparation of covalent bioconjugates can be found in ''Bioconjugate Techniques'', Hermanson, G.T. (1996), Academic Press Inc., which is hereby incorporated as reference (see below). It is preferred to make conservative substitutions in the polypeptide when the polypeptide has to be conjugated, as conservative substitutions secure that the impact of the substitution on the polypeptide structure is limited. In the case of providing additional amino groups this may be done by substitution of arginine to lysine, both residues being positively charged, but only the lysine having a free amino group suitable as an attachment groups. In the case of providing additional carboxylic acid groups the conservative substitution may for instance be an asparagine to aspartic acid or glutamine to glutamic acid substitution. These residues resemble each other in size and shape, except from the carboxylic groups being present on the acidic residues. In the case of providing SH-groups the conservative substitution may be done by changing threonine or serine to cysteine.

Chemical conjugation

For chemical conjugation, the protein variant needs to be incubate with an active or activated polymer and subsequently separated from the unreacted polymer. This can be done in solution followed by purification or it can conveniently be done using the immobilized protein variants, which can easily be exposed to different reaction environments and washes.

In the case were polymeric molecules are to be conjugated with the polypeptide in question and the polymeric molecules are not active they must be activated by the use of a suitable technique. It is also contemplated according to the invention to couple the polymeric molecules to the polypeptide through a linker. Suitable linkers are well-known to the skilled person. Methods and chemistry for activation of polymeric molecules as well as for conjugation of polypeptides are intensively described in the literature. Commonly used methods for activation of insoluble polymers include activation of functional groups with cyanogen bromide, periodate, glutaraldehyde, biepoxides, epichlorohydrin, divinylsulfone, carbodiimide, sulfonyl halides, trichlorotriazine etc. (see R.F. Taylor, (1991), "Protein immobilisation. Fundamental and applications", Marcel Dekker, N.Y.; S.S. Wong, (1992), "Chemistry of Protein Conjugation and Crosslinking", CRC Press, Boca Raton; G.T. Hermanson et al., (1993), "Immobilized Affinity Ligand Techniques", Academic Press, N.Y.). Some of the methods concern activation of insoluble polymers but are also applicable to activation of soluble polymers e.g. periodate, trichlorotriazine, sulfonylhalides, divinylsulfone, carbodiimide etc. The functional groups being amino, hydroxyl, thiol, carboxyl, aldehyde or sulfydryl on the polymer and the chosen attachment group on the protein must be considered in choosing the activation and conjugation chemistry which normally consist of i) activation of polymer, ii) conjugation, and iii) blocking of residual active groups.

In the following a number of suitable polymer activation methods will be described shortly. However, it is to be understood that also other methods may be used.

Coupling polymeric molecules to the free acid groups of polypeptides may be performed with the aid of diimide and for example amino-PEG or hydrazino-PEG (Pollak et al., (1976), J. Am. Chem. Soc., 98, 289-291) or diazoacetate/amide (Wong et al., (1992), "Chemistry of Protein Conjugation and Crosslinking", CRC Press).

Coupling polymeric molecules to hydroxy groups is generally very difficult as it must be performed in water. Usually hydrolysis predominates over reaction with hydroxyl groups.

Coupling polymeric molecules to free sulfhydryl groups can be achieved with special groups like maleimido or the orthopyridyl disulfide. Also vinylsulfone (US patent no. 5,414,135, (1995), Snow et al.) has a preference for sulfhydryl groups but is not as selective as the other mentioned.

Accessible arginine residues in the polypeptide chain may be targeted by groups comprising two vicinal carbonyl groups.

Techniques involving coupling of electrophilically activated PEGs to the amino groups of Lysines may also be useful. Many of the usual leaving groups for alcohols give rise to an amine linkage. For instance, alkyl sulfonates, such as tresylates (Nilsson et al., (1984), Methods in Enzymology vol. 104, Jacoby, W. B., Ed., Academic Press: Orlando, p. 56-66; Nilsson et al., (1987), Methods in Enzymology vol. 135; Mosbach, K., Ed.; Academic Press: Orlando, pp. 65-79; Scouten et al., (1987), Methods in Enzymology vol. 135, Mosbach, K., Ed., Academic Press: Orlando, 1987; pp 79-84; Crossland et al., (1971), J. Amr. Chem. Soc. 1971, 93, pp. 4217-4219), mesylates (Harris, (1985), supra; Harris et al., (1984), J. Polym. Sci.

Polym. Chem. Ed. 22, pp 341-352), aryl sulfonates like tosylates, and para-nitrobenzene sulfonates can be used.

Organic sulfonyl chlorides, e.g. Tresyl chloride, effectively converts hydroxy groups in a number of polymers, e.g. PEG, into good leaving groups (sulfonates) that, when reacted with nucleophiles like amino groups in polypeptides allow stable linkages to be formed between polymer and polypeptide. In addition to high conjugation yields, the reaction conditions are in general mild (neutral or slightly alkaline pH, to avoid denaturation and little or no disruption of activity), and satisfy the non-destructive requirements to the polypeptide.

Tosylate is more reactive than the mesylate but also less - stable decomposing into PEG, dioxane, and sulfonic acid (Zalipsky, (1995), Bioconjugate Chem., 6, 150-165). Epoxides may also been used for creating amine bonds but are much less reactive than the abovementioned groups.

Converting PEG into a chloroformate with phosgene gives rise to carbamate linkages to Lysines. Essentially the same reaction can be carried out in many variants substituting the chlorine with N-hydroxy succinimide (US patent no. 5,122,614, (1992); Zalipsky et al., (1992), Biotechnol. Appl. Biochem., 15, p. 100-114; Monfardini et al., (1995), Bioconjugate Chem., 6, 62-69, with imidazole (Allen et al., (1991), Carbohydr. Res., 213, pp 309-319), with para-nitrophenol, DMAP (EP 632 082 A1, (1993), Looze, Y.) etc. The derivatives are usually made by reacting the chloroformate with the desired leaving group. All these groups give rise to carbamate linkages to the peptide.

Furthermore, isocyanates and isothiocyanates may be employed, yielding ureas and thioureas, respectively.

Amides may be obtained from PEG acids using the same leaving groups as mentioned above and cyclic imid thrones (US patent no. 5,349,001, (1994), Greenwald et al.). The reactivity of these compounds are very high but may make the hydrolysis to fast.

PEG succinate made from reaction with succinic anhydride can also be used. The hereby comprised ester group make the conjugate much more susceptible to hydrolysis (US patent no. 5,122,614, (1992), Zalipsky). This group may be activated with N-hydroxy succinimide.

Furthermore, a special linker can be introduced. The most well studied being cyanuric chloride (Abuchowski et al., (1977), J. Biol. Chem., 252, 3578-3581; US patent no. 4,179,337, (1979), Davis et al.; Shafer et al., (1986), J. Polym. Sci. Polym. Chem. Ed., 24, 375-378.

Coupling of PEG to an aromatic amine followed by diazotation yields a very reactive diazonium salt, which can be reacted with a peptide in situ. An amide linkage may also be obtained by reacting an azlactone derivative of PEG (US patent no. 5,321,095, (1994), Greenwald, R. B.) thus introducing an additional amide linkage.

As some peptides do not comprise many Lysines it may be advantageous to attach more than one PEG to the same Lysine. This can be done e.g. by the use of 1,3-diamino-2-propanol.

PEGs may also be attached to the amino-groups of the enzyme with carbamate linkages (WO 95/11924, Greenwald et al.). Lysine residues may also be used as the backbone.
The coupling technique used in the examples is the N-succinimidyl carbonate conjugation technique descried in WO 90/13590 (Enzon).

In a preferred embodiment, the activated polymer is methyl-PEG which has been activated by N-succinimidyl carbonate as described WO 90/13590. The coupling can be carried out at alkaline conditions in high yields.

For coupling of polymers to the protein variants, it is preferred to use conditions similar to those described in WO96/17929 and WO99/00489 (Novo Nordisk A/S) e.g. mono or bis activated PEG´s of molecular weight ranging from 100 to 5000 Da. For instance, a methyl-PEG 350 could be activated with N-succinimidyl carbonate and incubated with protein variant at a molar ratio of more than 5 calculated as equivalents of activated PEG divided by moles of lysines in the protein of interest. For coupling to immobilized protein variant, the PEG:protein ratio should be optimized such that the PEG concentration is low enough for the buffer capacity to maintain alkaline pH throughout the reaction; while the PEG concentration is still high enough to ensure sufficient degree of modification of the protein. Further, it is important that the activated PEG is kept at conditions that prevent hydrolysis (i.e. dissolved in acid or solvents) and diluted directly into the alkaline reaction buffer. It is essential that primary amines are not present other than those occurring in the lysine residues of the protein. This can be secured by washing thoroughly in borate buffer. The reaction is stopped by sepa-

rating the fluid phase containing unreacted PEG from the solid phase containing protein and derivatized protein. Optionally, the solid phase can then be washed with tris buffer, to block any unreacted sites on PEG chains that might still be present.

Introduction of consensus sequences for post-translational modifications in the epitope areas

In another embodiment, the mutations are designed, such that recognition sites for post-translational modifications are introduced in the epitope areas, and the protein variant is expressed in a suitable host organism capable of the corresponding post-translational modification. These post-translational modifications may serve to shield the epitope and hence lower the immunogenicity of the protein variant relative to the protein backbone. Post-translational modifications include glycosylation, phosphorylation, N-terminal processing, acylation, ribosylation and sulfatation. A good example is N-glycosylation. N-glycosylation is found at sites of the sequence Asn-Xaa-Ser, Asn-Xaa-Thr, or Asn-Xaa-Cys, in which neither the Xaa residue nor the amino acid following the tripeptide consensus sequence is a proline (T. E. Creighton, 'Proteins - Structures and Molecular Properties, 2nd edition, W.H. Freeman and Co., New York, 1993, pp. 91-93). It is thus desirable to introduce such recognition sites in the sequence of the backbone protein. The specific nature of the glycosyl chain of the glycosylated protein variant may be linear or branched depending on the protein and the host cells. Another example is phosphorylation: The protein sequence can be modified so as to introduce serine phophorylation sites with the recognition sequence arg-arg-$(xaa)_n$-ser (where n = 0, 1, or 2),

which can be phosphorylated by the cAMP-dependent kinase or tyrosine phosphorylation sites with the recognition sequence - lys/arg - $(xaa)_3$ - asp/glu- $(xaa)_3$ - tyr, which can usually be phophorylated by tyrosine-specific kinases (T.E. Creighton, "Proteins- Structures and molecular properties", 2nd ed., Freeman, NY, 1993).

Randomized approaches to introduce modifications in epitope areas.

In order to generate protein variants, more than one amino acid residue may be substituted, added or deleted, these amino acids preferably being located in different epitope areas. In that case, it may be difficult to assess a priori how well the functionality of the protein is maintained while antigenicity is reduced, especially since the possible number of mutation-combinations becomes very large, even for a small number of mutations. In that case, it will be an advantage, to establish a library of diversified mutants each having one or more changed amino acids introduced and selecting those variants, which show good retention of function and at the same time a significant reduction in antigenicity.

A diversified library can be established by a range of techniques known to the person skilled in the art (Reetz MT; Jaeger KE, in 'Biocatalysis - from Discovery to Application' edited by Fessner WD, Vol. 200, pp. 31-57 (1999); Stemmer, Nature, vol. 370, p.389-391, 1994; Zhao and Arnold, Proc. Natl. Acad. Sci., USA, vol. 94, pp. 7997-8000, 1997; or Yano et al., Proc. Natl. Acad. Sci., USA, vol. 95, pp 5511-5515, 1998). These include, but are not limited to, 'spiked mutagenesis',

in which certain positions of the protein sequence are randomized by carring out PCR mutagenesis using one or more oligonucleotide primers which are synthesized using a mixture of nucleotides for certain positions (Lanio T, Jeltsch A, Biotechniques, Vol. 25(6), 958,962,964-965 (1998)). The mixtures of oligonucleotides used within each triplet can be designed such that the corresponding amino acid of the mutated gene product is randomized within some predetermined distribution function. Algorithms have been disclosed, which facilitate this design (Jensen LJ et al., Nucleic Acids Research, Vol. 26(3), 697-702 (1998)).

In an embodiment substitutions are found by a method comprising the following steps: 1) a range of substitutions, additions, and/or deletions are listed encompassing several epitope areas (preferably in the corresponding epitope sequences, anchor amino aids, and/or hot spots), 2) a library is designed which introduces a randomized subset of these changes in the amino acid sequence into the target gene, e.g. by spiked mutagenesis, 3) the library is expressed, and preferred variants are selected. In another embodiment, this method is supplemented with additional rounds of screening and/or family shuffling of hits from the first round of screening (J.E. Ness, et al, Nature Biotechnology, vol. 17, pp. 893-896, 1999) and/or combination with other methods of reducing immunogenicity by genetic means (such as that disclosed in WO92/10755).

The library may be designed, such that at least one amino acid of the epitope area is substituted. In a preferred embodiment at least one amino acid of the epitope sequence itself is changed, and in an even more preferred embodiment, one or more hot spot amino acids are changed. The library may be biased

such that towards introducing an amino acid of different size, hydrophilicity, and/or polarity relative to the original one of the 'protein backbone'. For example changing a small amino acid to a large amino acid, a hydrophilic amino acid to a hydrophobic amino acid, a polar amino acid to a non-polar amino acid or a basic to an acidic amino acid. Other changes may be the addition or deletion of at least one amino acid of the epitope area, preferably deleting an anchor amino acid. Furthermore, substituting some amino acids and deleting or adding others may change an epitope.

Diversity in the protein variant library can be generated at the DNA triplet level, such that individual codons are variegated e.g. by using primers of partially randomized sequence for a PCR reaction. Further, several techniques have been described, by which one can create a library with such diversity at several locations in the gene, which are too far apart to be covered by a single (spiked) oligonucleotide primer. These techniques include the use of in vivo recombination of the individually diversified gene segments as described in WO 97/07205 on page 3, line 8 to 29 or by using DNA shuffling techniques to create a library of full length genes that combine several gene segments each of which are diversified e.g. by spiked mutagenesis (Stemmer, Nature 370, pp. 389-391, 1994 and US 5,605,793 and 5,830,721). In the latter case, one can use the gene encoding the "protein backbone" as a template double-stranded polynucleotide and combining this with one or more single or double-stranded oligonucleotides as described in claim 1 of US 5,830,721. The single- stranded oligonucleotides could be partially randomized during synthesis. The double- stranded oligonucleotides could be PCR products incorporating diversity in a specific region. In both cases, one can

dilute the diversity with corresponding segments containing the sequence of the backbone protein in order to limit the number of changes that are on average introduced. As mentioned above, methods have been established for designing the ratios of nucleotides (A; C; T; G) used at a particular codon during primer synthesis, so as to approximate a desired frequency distribution among a set of desired amino acids at that particular codon. This allows one to bias the partially randomized mutagenesis towards e.g. introduction of post-translational modification sites, chemical modification sites, or simply amino acids that are different from those that define the epitope or the epitope area. One could also approximate a sequence in a given location or epitope area to the corresponding location on a homologous, human protein.

Occasionally, one would be interested in testing a library that combines a number of known mutations in different locations in the primary sequence of the 'protein backbone'. These could be introduced post-translational or chemical modification sites, or they could be mutations, which by themselves had proven beneficial for one reason or another (e.g. decreasing antigenicity, or improving specific activity, performance, stability, or other characteristics). In such cases, it may be desirable to create a library of diverse combinations of known sequences. For example if 12 individual mutations are known, one could combine (at least) 12 segments of the 'protein backbone' gene in which each segment is present in two forms: one with and one without the desired mutation. By varying the relative amounts of those segments, one could design a library (of size 2^{12}) for which the average number of mutations per gene can be predicted. This can be a useful way of combining elements that by themselves give some, but not sufficient ef-

fect, without resorting to very large libraries, as is often the case when using 'spiked mutagenesis'. Another way to combine these 'known mutations' could be by using family shuffling of oligomeric DNA encoding the known changes with fragments of the full length wild type sequence.

Assays for reduced allergenicity

When protein variants have been constructed based on the methods described in this invention, it is desirable to confirm their antibody binding capacity, functionality, immunogenicity and/or allergenicity using a purified preparation. For that use, the protein variant of interest can be expressed in larger scale, purified by conventional techniques, and the antibody binding and functionality should be examined in detail using dose-response curves and e.g. direct or competitive ELISA (C-ELISA).

The potentially reduced allergenicity (which is likely, but not necessarily true for a variant w. low antibody binding) should be tested in in vivo or in vitro model systems: e.g. an in vitro assays for immunogenicity such as assays based on cytokine expression profiles or other proliferation or differentiation responses of epithelial and other cells incl. B-cells and T-cells. Further, animal models for testing allergenicity should be set up to test a limited number of protein variants that show desired characteristics in vitro. Useful animal models include the guinea pig intratracheal model (GPIT) (Ritz, et al. Fund. Appl. Toxicol., 21, pp. 31-37, 1993), mouse subcutaneous (mouse-SC) (WO 98/30682, Novo Nordisk), the rat intratracheal (rat-IT) (WO 96/17929, Novo Nord-

isk), and the mouse intranasal (MINT) (Robinson et al., Fund. Appl. Toxicol. 34, pp. 15-24, 1996) models.

The immunogenicity of the protein variant is measured in animal tests, wherein the animals are immunised with the protein variant and the immune response is measured. Specifically, it is of interest to determine the allergenicity of the protein variants by repeatedly exposing the animals to the protein variant by the intratracheal route and following the specific IgG and IgE titers. Alternatively, the mouse intranasal (MINT) test can be used to assess the allergenicity of protein variants. By the present invention the allergenicity is reduced at least 3 times as compared to the allergenicity of the parent protein, preferably 10 times reduced, more preferably 50 times.

However, the present inventors have demonstrated that the performance in ELISA correlates closely to the immunogenic responses measured in animal tests. To obtain a useful reduction of the allergenicity of a protein, the IgE binding capacity of the protein variant must be reduced to at least below 75 %, preferably below 50 %, more preferably below 25 % of the IgE binding capacity of the parent protein as measured by the performance in IgE ELISA, given the value for the IgE binding capacity of the parent protein is set to 100 %.

Thus a first asessment of the immunogenicity and/or allergenicity of a protein can be made by measuring the antibody binding capacity or antigenicity of the protein variant using appropriate antibodies. This approach has also been used in the literature (WO 99/47680).

Assays for altered immunotherapeutic effect

The immunotherapeutic effect of allergen vaccines can be assessed a number of different ways. One is to measure the specific IgE binding, the reduction of which indicates a better allergen vaccine potential (WO 99/47680, ALK-ABELLÓ). Also, several cellular assays could be employed to show the modified immuneresponse indicative of good allergen vaccine potential as shown in several publications, all of which are hereby incorporated by reference (van Neerven et al, '' T lymphocyte responses to allergens: Epitope-specificity and clinical relevance'', Immunol Today, 1996, vol. 17, pp. 526-532; Hoffmann et al., Allergy, 1999, vol. 54, pp. 446-454, WO99/07880).

Eventually, clinical trials with allergic patients could be employed using cellular or clinical end-point measurements. (Ebner et al., Clin. Exp. All., 1997, vol. 27, pp. 107-1015; Int. Arch. Allergy Immunol., 1999, vol. 119, pp 1-5).

Determining functionality

A wide variety of protein functionality assays are available in the literature. Especially, those suitable for automated analysis are useful for this invention. Several have been published in the literature such as protease assays (WO99/34011, Genencor International; J.E. Ness, et al, Nature Biotechn., 17, pp. 893-896, 1999), oxidoreductase assays (Cherry et al., Nature Biotechn., 17 , pp. 379-384, 1999, and assays for several other enzymes (WO99/45143, Novo Nordisk). Those assays that employ soluble substrates can be employed

for direct analysis of functionality of immobilized protein variants.

Cross-reactivity

A related objective is to reduce cross-reactivity between 'commercial allergens' and 'environmental allergens'. Cross-reactivities between food allergens of different origin are well-known (Akkerdaas et al, Allergy 50, pp 215-220, 1995). Similarly, cross-reactivities between other environmental allergens (like pollen, dust mites etc.) and commercial allergens (like enzyme proteins) have been established in the literature (J. All. Clin. Immunol., 1998, vol. 102, pp. 679-686 and by the present inventors. The molecular reason for this cross-reactivity can be explored using epitope mapping. By finding epitope patterns using antibodies raised against environmental allergen (donor protein) and mapping this information on a commercial allergen (the acceptor protein), one may find the epitopes that are common to both proteins, and hence responsible for the cross-reactivity. Obviously, one can also use the commercial allergen as donor and the environmental allergen as acceptor. By modifying the commercial allergen using protein engineering in the epitope areas identified as described above, one can reduce the cross-reactivity of the commercial allergen variant towards the environmental allergens (and vice versa). Hence, the use of the modified commercial allergens would be safer than using the unmodified commercial allergen.

Testing of this approach would be done using an antibody-binding assay with the protein variant (and its parent protein

as control) and antibodies raised against the protein that cross-reacts with the parent protein. The method is otherwise identical to those described in the Methods section for characterization of allergencitiy and antigenicity.

Commercial enzyme applications

Industrial applications

Another aspect of the invention is a composition comprising at least one protein (polypeptide) or enzyme of the invention. The composition may comprise other polypeptides, proteins or enzymes and/or ingredients normally used in personal care products, such as shampoo, soap bars, skin lotion, skin creme, hair dye, toothpaste, household articles, agro chemicals, personal care products, such as cleaning preparations e.g. for contact lenses, cosmetics, toiletries, oral and dermal pharmaceuticals, compositions used for treating textiles, compositions used for manufacturing food, e.g. baking, and feed etc.

Examples of said proteins(polypeptides)/enzymes include enzymes exhibiting protease, lipase, oxidoreductase, carbohydrase, transferase, such as transglutaminase, phytase and/or antimicrobial polypeptide activity. These enzymes may be present as conjugates with reduced activity.

The protein of the invention may furthermore typically be used in detergent composition. It may be included in the detergent composition in the form of a non-dusting granulate, a stabilized liquid, or a protected enzyme. Non-dusting granulates may be produced, e.g., as disclosed in US 4,106,991 and 4,661,452 (both

to Novo Industri A/S) and may optionally be coated by methods known in the art. Examples of waxy coating materials are poly(ethylene oxide) products (polyethylene glycol, PEG) with mean molecular weights of 1000 to 20000; ethoxylated nonylphenols having from 16 to 50 ethylene oxide units; ethoxylated fatty alcohols in which the alcohol contains from 12 to 20 carbon atoms and in which there are 15 to 80 ethylene oxide units; fatty alcohols; fatty acids; and mono- and di- and triglycerides of fatty acids. Examples of film-forming coating materials suitable for application by fluid bed techniques are given in patent GB 1483591. Liquid enzyme preparations may, for instance, be stabilized by adding a polyol such as propylene glycol, a sugar or sugar alcohol, lactic acid or boric acid according to established methods. Other enzyme stabilizers are well known in the art. Protected enzymes may be prepared according to the method disclosed in EP 238,216.

The detergent composition may be in any convenient form, e.g. as powder, granules, paste or liquid. A liquid detergent may be aqueous, typically containing up to 70% water and 0-30% organic solvent, or non-aqueous.

The detergent composition comprises one or more surfactants, each of which may be anionic, nonionic, cationic, or zwitterionic. The detergent will usually contain 0-50% of anionic surfactant such as linear alkylbenzenesulfonate (LAS), alpha-olefinsulfonate (AOS), alkyl sulfate (fatty alcohol sulfate) (AS), alcohol ethoxysulfate (AEOS or AES), secondary alkanesulfonates (SAS), alpha-sulfo fatty acid methyl esters, alkyl- or alkenylsuccinic acid, or soap. It may also contain 0-40% of nonionic surfactant such as alcohol ethoxylate (AEO or AE), carboxylated alcohol ethoxylates, nonylphenol ethoxylate, alkyl-

polyglycoside, alkyldimethylamine oxide, ethoxylated fatty acid monoethanolamide, fatty acid monoethanolamide, or polyhydroxy alkyl fatty acid amide (e.g. as described in WO 92/06154).

The detergent composition may additionally comprise one or more other enzymes, such as e.g. proteases, amylases, lipases, cutinases, cellulases, peroxidases, oxidases, and further antimicrobial polypeptides.

The detergent may contain 1-65% of a detergent builder or complexing agent such as zeolite, diphosphate, triphosphate, phosphonate, citrate, nitrilotriacetic acid (NTA), ethylenediaminetetraacetic acid (EDTA), diethylenetriaminepentaacetic acid (DTMPA), alkyl- or alkenylsuccinic acid, soluble silicates or layered silicates (e.g. SKS-6 from Hoechst). The detergent may also be unbuilt, i.e. essentially free of detergent builder.

The detergent may comprise one or more polymers. Examples are carboxymethylcellulose (CMC), poly(vinylpyrrolidone) (PVP), polyethyleneglycol (PEG), poly(vinyl alcohol) (PVA), polycarboxylates such as polyacrylates, maleic/acrylic acid copolymers and lauryl methacrylate/acrylic acid copolymers.

The detergent may contain a bleaching system which may comprise a H_2O_2 source such as perborate or percarbonate which may be combined with a peracid-forming bleach activator such as tetraacetylethylenediamine (TAED) or nonanoyloxybenzenesulfon-ate (NOBS). Alternatively, the bleaching system may comprise peroxyacids of, e.g., the amide, imide, or sulfone type.

The detergent composition of the invention comprising the polypeptide of the invention may be stabilized using conventional

stabilizing agents, e.g. a polyol such as propylene glycol or glycerol, a sugar or sugar alcohol, lactic acid, boric acid, or a boric acid derivative such as, e.g., an aromatic borate ester, and the composition may be formulated as described in, e.g., WO 92/19709 and WO 92/19708.

The detergent may also contain other conventional detergent ingredients such as, e.g., fabric conditioners including clays, foam boosters, suds suppressors, anti-corrosion agents, soil-suspending agents, anti-soil-redeposition agents, dyes, bactericides, optical brighteners, or perfume.

The pH (measured in aqueous solution at use concentration) will usually be neutral or alkaline, e.g. in the range of 7-11.

Dishwashing composition

Further, a modified enzyme according to the invention may also be used in dishwashing detergents.

Dishwashing detergent compositions comprise a surfactant which may be anionic, non-ionic, cationic, amphoteric or a mixture of these types. The detergent will contain 0-90% of non-ionic surfactant such as low- to non-foaming ethoxylated propoxylated straight-chain alcohols.

The detergent composition may contain detergent builder salts of inorganic and/or organic types. The detergent builders may be subdivided into phosphorus-containing and non-phosphorus-containing types. The detergent composition usually contains 1-90% of detergent builders.

Examples of phosphorus-containing inorganic alkaline detergent builders, when present, include the water-soluble salts especially alkali metal pyrophosphates, orthophosphates, and polyphosphates. An example of phosphorus-containing organic alkaline detergent builder, when present, includes the water-soluble salts of phosphonates. Examples of non-phosphorus-containing inorganic builders, when present, include water-soluble alkali metal carbonates, borates and silicates as well as the various types of water-insoluble crystalline or amorphous alumino silicates of which zeolites are the best-known representatives.

Examples of suitable organic builders include the alkali metal, ammonium and substituted ammonium, citrates, succinates, malonates, fatty acid sulphonates, carboxymetoxy succinates, ammonium polyacetates, carboxylates, polycarboxylates, aminopolycarboxylates, polyacetyl carboxylates and polyhydroxsulphonates.

Other suitable organic builders include the higher molecular weight polymers and co-polymers known to have builder properties, for example appropriate polyacrylic acid, polymaleic and polyacrylic/polymaleic acid copolymers and their salts.

The dishwashing detergent composition may contain bleaching agents of the chlorine/bromine-type or the oxygen-type. Examples of inorganic chlorine/bromine-type bleaches are lithium, sodium or calcium hypochlorite and hypobromite as well as chlorinated trisodium phosphate. Examples of organic chlorine/bromine-type bleaches are heterocyclic N-bromo and N-chloro imides such as trichloroisocyanuric, tribromoisocyanuric, dibromoisocyanuric and dichloroisocyanuric acids, and salts thereof with water-

solubilizing cations such as potassium and sodium. Hydantoin compounds are also suitable.

The oxygen bleaches are preferred, for example in the form of an inorganic persalt, preferably with a bleach precursor or as a peroxy acid compound. Typical examples of suitable peroxy bleach compounds are alkali metal perborates, both tetrahydrates and monohydrates, alkali metal percarbonates, persilicates and perphosphates. Preferred activator materials are TAED and glycerol triacetate.

The dishwashing detergent composition of the invention may be stabilized using conventional stabilizing agents for the enzyme(s), e.g. a polyol such as e.g. propylene glycol, a sugar or a sugar alcohol, lactic acid, boric acid, or a boric acid derivative, e.g. an aromatic borate ester.

The dishwashing detergent composition of the invention may also contain other conventional detergent ingredients, e.g. deflocculant material, filler material, foam depressors, anti-corrosion agents, soil-suspending agents, sequestering agents, anti-soil redeposition agents, dehydrating agents, dyes, bactericides, fluorescers, thickeners and perfumes.

Finally, the enzyme of the invention may be used in conventional dishwashing-detergents, e.g. in any of the detergents described in any of the following patent publications:

EP 518719, EP 518720, EP 518721, EP 516553, EP 516554, EP 516555, GB 2200132, DE 3741617, DE 3727911, DE 4212166,

DE 4137470, DE 3833047, WO 93/17089, DE 4205071, WO 52/09680, WO 93/18129, WO 93/04153, WO 92/06157, WO 92/08777, EP 429124, WO 93/21299, US 5141664, EP 561452, EP 561446, GB 2234980, WO 93/03129, EP 481547, EP 530870, EP 533239, EP 554943, EP 346137, US 5112518, EP 318204, EP 318279, EP 271155, EP 271156, EP 346136, GB 2228945, CA 2006687, WO 93/25651, EP 530635, EP 414197, US 5240632.

Personal care applications

A particularly useful application area for low allergenic proteins or of proteins with low cross-reactivity to environmental allergens would be in personal care products where the end-user is in close contact with the protein, and where certain problems with allergenicity has been encountered in experimental set-ups (Kelling et al., J. All. Clin. Imm., 1998, Vol. 101, pp. 179-187 and Johnston et al., Hum. Exp. Toxicol., 1999, Vol.18, p. 527).

First of all the conjugate or compositions of the invention can advantageously be used for personal care products, such as hair care and hair treatment products. This include products such as shampoo, balsam, hair conditioners, hair waving compositions, hair dyeing compositions, hair tonic, hair liquid, hair cream, shampoo, hair rinse, hair spray.

Further contemplated are oral care products such as dentifrice, oral washes, chewing gum.

Also contemplated are skin care products and cosmetics, such as skin cream, skin milk, cleansing cream, cleansing lotion, cleansing milk, cold cream, cream soap, nourishing essence, skin lotion, milky lotion, calamine lotion, hand cream, powder soap, transparent soap, sun oil, sun screen, shaving foam, shaving cream, baby oil lipstick, lip cream, creamy foundation, face powder, powder eye-shadow, powder, foundation, make-up base, essence powder, whitening powder.

Also for contact lenses hygiene products the conjugate of the invention can be used advantageously. Such products include cleaning and disinfection products for contact lenses.

Proteases

Proteases are well-known active ingredients for cleaning of contact lenses. They hydrolyse the proteinaceous soil on the lens and thereby makes it soluble. Removal of the protein soil is essential for the wearing comfort.

Proteases are also effective ingredients in skin cleaning products, where they remove the upper layer of dead keratinaseous skin cells and thereby make the skin look brighter and fresher.

Proteases are also used in oral care products, especially for cleaning of dentures, but also in dentifrices.

Further, proteases are used in toiletries, bath and shower products, including shampoos, conditioners, lotions, creams, soap bars, toilet soaps, and liquid soaps.

Lipases

Lipases can be applied for cosmetic use as active ingredients in skin cleaning products and anti-acne products for removal of excessive skin lipids, and in bath and shower products such as creams and lotions as active ingredients for skin care.

Lipases can also be used in hair cleaning products (e.g. shampoos) for effective removal of sebum and other fatty material from the surface of hair.

Lipases are also effective ingredients in products for cleaning of contact lenses, where they remove lipid deposits from the lens surface.

Oxidoreductases

The most common oxidoreductase for personal care purposes is an oxidase (usually glucose oxidase) with substrate (e.g. glucose) that ensures production of H_2O_2, which then will initiate the oxidation of for instance SCN^- or I^- into antimicrobial reagents ($SCNO^-$ or I_2) by a peroxidase (usually lactoperoxidase). This enzymatic complex is known in nature from e.g. milk and saliva.

It is being utilised commercially as anti-microbial system in oral care products (mouth rinse, dentifrice, chewing gum) where it also can be combined with an amyloglucosidase to produce the glucose. These systems are also known in cosmetic products for preservation.

Anti-microbial systems comprising the combination of an oxidase and a peroxidase are know in the cleaning of contact lenses.

Another application of oxidoreductases is oxidative hair dyeing using oxidases, peroxidases and laccases.

Free radicals formed on the surface of the skin (and hair) known to be associated with the ageing process of the skin (spoilage of the hair). The free radicals activate chain reactions that lead to destruction of fatty membranes, collagen, and cells. The application of free radical scavengers such as Superoxide dismutase into cosmetics is well known (R. L. Goldemberg, DCI, Nov. 93, p. 48-52).

Protein disulfide isomerase (PDI) is also an oxidoreductase. It can be utilised for waving of hair (reduction and reoxidation of disulfide bonds in hair) and repair of spoiled hair (where the damage is mainly reduction of existing disulfide bonds).

Carbohydrases

Plaque formed on the surface of teeth is composed mainly of polysaccharides. They stick to the surface of the teeth and the microorganisms. The polysaccharides are mainly α-1,6 bound glucose (dextran) and α-1,3 bound glucose (mutan). The application of different types of glucanases such as mutanase and dextranase helps hydrolysing the sticky matrix of plaque, making it easier to remove by mechanical action.

Also other kinds of biofilm for instance the biofilm formed in lens cases can be removed by the action of glucanases.

Food and Feed

Further conjugated enzymes or polypeptides with reduced immunogenicity according to the invention may advantageously be used in the manufacturing of food and feed.

Proteases

The gluten in wheat flour is the essential ingredient responsible for the ability of flour to be used in baked foodstuffs. Proteolytic enzymes are sometimes needed to modify the gluten phase of the dough, e.g. a hard wheat flour can be softened with a protease.

Neutrase® is a commercially available neutral metallo protease that can be used to ensure a uniform dough quality and bread texture, and to improve flavour. The gluten proteins are degraded either moderately or more extensively to peptides, whereby close control is necessary in order to avoid excessive softening of the dough.

Proteases are also used for modifying milk protein.

To coagulate casein in milk when producing cheese proteases such as rennet or chymosin may be used.

In the brewery industry proteases are used for brewing with unmalted cereals and for controlling the nitrogen content.

In animal feed products proteases are used so to speak to expand the animals digestion system.

Lipases

The application of lipase in the baking industry is rather new. Addition of lipase results in improved dough properties and an improved breadmaking quality in terms of larger volume, improved crumb structure and whiter crumb colour. The observed effect can be explained by a mechanism where the lipase changes the interaction between gluten and some lipids fragment during dough mixing. This results in an improved gluten network.

The flavour development of blue roan cheese (e.g. Danablue), certain Italian type cheese, and other dairy products containing butter-fat, are dependent on the degradation of milk fat into free fatty acids. Lipases may be used for developing flavour in such products.

In the oil- and fat producing industry lipases are used e.g. to minimize the amount of undesirable side-products, to modify fats by interesterification, and to synthesis of esters.

Oxidoreductases

Further oxidoreductases with reduced immunogenicity according to the invention may advantageously be used in the manufacturing of food and feed.

Several oxidoreductases are used for baking, glucose oxidase, lipoxygenase, peroxidase, catalase and combinations hereof. Traditionally, bakers strengthen gluten by adding ascorbic acid and potassium bromate. Some oxidoreductases can be used to replace bromate in dough systems by oxidation of free sulfydryl units in

gluten proteins. Hereby disulphide linkages are formed resulting in stronger, more elastic doughs with greater resistance.

Gluzyme™ (Novo Nordisk A/S) is a glucose oxidase preparation with catalase activity that can be used to replace bromate. The dough strengthen is measured as greater resistance to mechanical shock, better oven spring and larger loaf volume.

Carbohydrases

Flour has varying content of amylases leading to differences in the baking quality. Addition of amylases can be necessary in order to standardize the flour. Amylases and pentosanases generally provide sugar for the yeast fermentation, improve the bread volume, retard retrogradation, and decrease the staling rate and stickiness that results from pentosan gums. Examples of carbohydrases are given below.

Certain maltogenic amylases can be used for prolonging the shelf life of bread for two or more days without causing gumminess in the product. Selectively modifies the gelatinized starch by cleaving from the non-reducing end of the starch molecules, low molecular wight sugars and dextrins. The starch is modified in such a way that retrogradation is less likely to occur. The produced low-molecular-weight sugars improve the baked goods water retention capacity without creating the intermediate-length dextrins that result in gumminess in the finished product. The enzyme is inactivated during bread baking, so it can be considered a processing aid that does not have to be declared on the label. Overdosing of Novamyl can almost be excluded.

The bread volume can be improved by fungal α-amylases which further provide good and uniform structure of the bread crumb. Said α-amylases are endoenzymes that produce maltose, dextrins and glucose. Cereal and some bacterial α-amylases are inactivated at temperatures above the gelatinization temperature of starch, therefore when added to wheat dough it results in a low bread volume and a sticky bread interior. Fungamyl has the advantage of being thermolabile and is inactivated just below the gelatinization temperature.

Enzyme preparations containing a number of pentosanase and hemicellulase activities can improve the handling and stability of the dough, and improves the freshness, the crumb structure and the volume of the bread.

By hydrolysing the pentosans fraction in flour, it will lose a great deal of its water-binding capacity, and the water will then be available for starch and gluten. The gluten becomes more pliable and extensible, and the starch gelatinizes more easily. Pentosanases can be used in combination with or as an alternative to emulsifiers.

Further carbohydrases are user for producing syrups from starch, which are widely used in soft drinks, sweets, meat products, dairy products, bread products, ice cream, baby food, jam etc.

The conversion of starch is normally carried out three steps. First the starch is liquefied, by the use of α-amylases. Maltodextrins, primary consisting of oligosaccharides and dextrins, are obtained.
The mixture is then treated with an amyloglucosidase for hydrolysing the oligosaccharides and dextrins into glucose. This way

a sweeter product is obtained. If high maltose syrups are desired β-amylases alone or in combination with a pullulanase (debranching enzyme) may be used.

The glucose mixture can be made even sweeter by isomerization to fructose. For this an immobilized glucose isomerase can be used.

In the sugar industry, it is common practice to speed up the break down of present starch in cane juices. Thereby the starch content in the raw sugar is reduced and filtration at the refinery facilitated.
Furthermore dextranases are used to break down dextran in raw sugar juices and syrups.

In the alcohol industry α-amylases is advantageously being used for thinning of starch in distilling mashes.

In the brewing industry α-amylases is used for adjunct liquefaction.

In the dairy industry β-galactosidases (lactase) is used when producing low lactose milk for persons suffering from lactose malabsorption.

When flavoured milk drinks are produced from lactase-treated milk, the addition of sugar can be reduced without reducing the sweetness of the product.

In the production of condensed milk, lactose crystallization can be avoided by lactase treatment, and the risk of thickening caused by casein coagulation in lactose crystals is thus reduced.

When producing ice cream made from lactase-treated milk (or whey) no lactose crystals will be formed and the defect, sandiness, will not occur.

Further, xylanases are known to be used within a number of food/feed industrial applications as described in WO 94/21785 (Novo Nordisk A/S).

α-amylases are used in the animal feed industry to be added to cereal-containing feed to improve the digestibility of starch.

Anti-microbial polypeptides

Certain bacteriolytic enzymes may be used e.g. to wash carcasses in the meat packing industry (see US patent no. 5,354,681 from Novo Industri A/S)

Transferases

Transglutaminases with reduced immunogenicity according to the invention may advantageously be used in the manufacturing of food and feed.

Transglutaminases has the ability to crosslinking protein.

This property can be used for gelling of aqueous phases containing proteins. This may be used for when producing of spreads (DK patent application no. 1071/84 from Novo Nordisk A/S).

Transglutaminases are being used for improvement of baking quality of flour e.g. by modifying wheat flour to be used in the preparation of cakes with improved properties, such as improved taste, dent, mouth-feel and a higher volume (see JP 1-110147).

Further producing paste type food material e.g. used as fat substitution in foods as ice cream, toppings, frozen desserts, mayonnaises and low fat spreads (see WO 93/22930 from Novo Nordisk A/S).

Furthermore for preparation of gels for yoghurt, mousses, cheese, puddings, orange juice, from milk and milk-like products, and binding of chopped meat product, improvement of taste and texture of food proteins (see WO 94/21120 and WO 94/21129 from Novo Nordisk A/S).

Phytases

Phytases of the invention may advantageously be used in the manufacturing of food, such as breakfast cereal, cake, sweets, drinks, bread or soup etc., and animal feed.

Phytases may be used either for exploiting the phosphorus bound in the phytate/phytic acid present in vegetable protein sources or for exploiting the nutritionally important minerals bound in phytic acid complexes.

Microbial phytase may be added to feedstuff of monogastric animals in order to avoid supplementing the feed with inorganic phosphorus (see US patent no. 3,297,548).

Further phytases may be used in soy processing. Soyabean meal may contain high levels of the anti-nutritional factor phytate which renders this protein source unsuitable for application in baby food and feed for fish, calves and other non-ruminants, since the phytate chelates essential minerals present therein (see EP 0 420 358).

Also for baking purposes phytases may be used. Bread with better quality can be prepared by baking divided pieces of a dough containing wheat flour etc. and phytase (see JP-0-3076529-A).

A high phytase activity as in koji mold are known to be used for producing refined sake (see JP-0-6070749-A).

Textile applications

Proteases

Proteases are used for degumming and sand washing of silk.

Lipases

Lipases are used for removing fatty matter containing hydrophobic esters (e.g. triglycerides) during the finishing of textiles (see e.g. WO 93/13256 from Novo Nordisk A/S).

Oxidoreductases

In bleach clean up of textiles catalases may serve to remove excess hydrogen peroxide.

Carbohydrases

Cellulolytic enzymes are widely used in the finishing of denim garments in order to provide a localized variation in the colour density of the fabric (Enzyme facilitated "stone wash").

Also cellulolytic enzymes find use in the bio-polishing process. Bio-Polishing is a specific treatment of the yarn surface which improves fabric quality with respect to handle and appearance without loss of fabric wettability. Bio-polishing may be obtained by applying the method described e.g. in WO 93/20278.

During the weaving of textiles, the threads are exposed to considerable mechanical strain. In order to prevent breaking, the threads are usually reinforced by the coating (sizing) with a gelatinous substance (size). The most common sizing agent is starch in native or modified form. A uniform and durable finish can thus be obtained only after removal of the size from the fabric, the so-called desizing. Desizing of fabrics sized with a size containing starch or modified starch is preferably facilitated by use of amylolytic enzymes.

Oral and dermal pharmaceuticals

Proteases

Different combinations of highly purified proteases (e.g. Trypsin and Chymotrypsin) are used in pharmaceuticals to be taken orally, and dermal pharmaceuticals for combating e.g inflammations, edemata and injuries.

Leather production

Transferase

Transglutaminase is known to be used to casein-finishing leather by acting as a hardening agent (see WO 94/13839 from Novo Nordisk).

Hard surface cleaning

Cleaning of hard surfaces e.g. in the food industry is often difficult, as equipment used for producing dairies, meat, sea food products, beverages etc. often have a complicated shape. The use of surfactant compositions in the form gels and foams comprising enzymes have shown to facilitate and improve hard surface cleaning. Enzymes, which advantageously may be added in such surfactant compositions, are in particular proteases, lipases, amylases and cellulases.

Such hard surface cleaning compositions comprising enzymes may also advantageously be used in the transport sector, for instance for washing cars and for general vessel wash.

Furthermore this invention relates to the method by which the protein variants are being synthesised and expressed in host cells. This is achieved by culturing host cells capable of expressing a polypeptide in a suitable culture medium to obtain expression and secretion of the polypeptide into the medium, followed by isolation of the polypeptide from the culture medium. The host cell may be any cell suitable for the large-

scale production of proteins, capable of expressing a protein and being transformed by an expression vector.

The host cell comprises a DNA construct as defined above, optionally the cells may be transformed with an expression vector comprising a DNA construct as defined above. The host cell is selected from any suitable cell, such as a bacterial cell, a fungal cell, an animal cell, such as an insect cell or a mammalian cell, or a plant cell.

Immunotherapy

A number of vaccination approaches have been described to for infective diseases as well as for non-infective diseases (such as cancers). In a number of cases, the antigen provided is an isolated protein or protein-adjuvant mixture and more and more often, the protein is recombinant (e.g. the hepatitits B vaccine from Merck & Co). In these cases, it could be desirable to modify the immunogenicity of the antigen vaccine, such that it offers a stronger or more specific protection. This can be achieved by protein engineering of the amino acid sequence of the antigen, and would be greatly facilitated by the use of the methods of this invention for identification of epitopes on the antigen vaccine to be the favored sites for modification.

There are several examples of vaccine molecules that have been engineered to achieve a specific immune protection against virus, parasites or cancer (Ryu and Nam, Biotechnol. Prog., 2000, vol. 16 pp.2-16; and references cited therein). "The goal is often to vaccinate with a minimal strucutre consisting

of a well-defined antigen, to stimulate an effective specific immune response, while avoiding potentially hazardous risks'' (Ryu and Nam, Biotechnol. Prog., 2000, vol. 16 pp.2-16). Thus, the methods of this invention can be used to identify such minimal structures that define an antigen (or epitope thereof) whether in the form of the parent protein scaffold with a number of mutations introduced in it, or whether it is in the form of the antibody binding peptides themselves.

Allergen vaccines

Today, a patient suffering allergic disease may be subjected to allergy vaccine therapy using allergens selected on the basis of testing the specificity of the patient's serum IgE against a bank of allergen extracts (or similar specificity tests of the patient's sensibilization such as skin prick test.

One could improve the quality of characterization by using antibody binding peptides corresponding to various epitope sequences on the protein allergens of interest. This would require a kit comprising reagents for such specificity characterization, e.g. the antibody binding peptides of desired specificity. It would be preferred to use antibody binding sequences in the kit, which correspond to defined epitope sequences known to be specific for the allergen under investigation (i.e. not identified on other allergens and/or not cross-reacting with sera raised against other allergens). This kit would be useful to specifying which allergy the patient is suffering from. This kit will lead to a more specific answer

than those kits used today, and hence to a better selection of allergen vaccine therapy for the individual patient.

Further, the knowledge about cross-reacting epitopes may improve vaccine development.

In an extension of this approach, one could also characterize the patient's serum by identifying the corresponding antibody binding peptides among a random display library using the aforementioned methods. This again may lead to a better selection of allergen vaccine therapy.

Further, one could use the individual antibody binding sequences as allergen vaccines leading to more specific allergen vaccine. These antibody binding sequences could be administered in an isolated form or fused to a membrane protein of the phage display system, or to another protein, which may have beneficial effect for the immunoprotective effect of the antibody binding peptide (Dalum et al., Nature Biotechnology, 1999, Vol. 17, pp. 666-669).

D) Variations possible.

Parent protein

The "parent protein" can in principle be any protein molecule of biological origin, non-limiting examples of which are peptides, polypeptides, proteins, enzymes, post-translationally

modified polypeptides such as lipopeptides or glycosylated peptides, anti-microbial peptides or molecules, and proteins having pharmaceutical properties etc.

Accordingly the invention relates to a method, wherein the ''parent protein'' is chosen from the group consisting of polypeptides, small peptides, lipopeptides, antimicrobials, and pharmaceutical polypeptides.

The term "pharmaceutical polypeptides" is defined as polypeptides, including peptides, such as peptide hormones, proteins and/or enzymes, being physiologically active when introduced into the circulatory system of the body of humans and/or animals.

Pharmaceutical polypeptides are potentially immunogenic as they are introduced into the circulatory system.

Examples of ''pharmaceutical polypeptides'' contemplated according to the invention include insulin, ACTH, glucagon, somatostatin, somatotropin, thymosin, parathyroid hormone, pigmentary hormones, somatomedin, erythropoietin, luteinizing hormone, chorionic gonadotropin, hypothalmic releasing factors, antidiuretic hormones, thyroid stimulating hormone, relaxin, interferon, thrombopoietin (TPO) and prolactin.

However, the proteins are preferably to be used in industry, housekeeping and/or medicine, such as proteins used in personal care products (for example shampoo; soap; skin, hand and face lotions; skin, hand and face cremes; hair dyes; toothpaste), food (for example in the baking industry), detergents and pharmaceuticals.

Antimicrobial peptides.

The antimicrobial peptide (AMP) may be, e.g., a membrane-active antimicrobial peptide, or an antimicrobial peptide affecting/interacting with intracellular targets, e.g. binding to cell DNA. The AMP is generally a relatively short peptide, consisting of less than 100 amino acid residues, typically 20-80 residues. The antimicrobial peptide has bactericidal and/or fungicidal effect, and it may also have antiviral or antitumour effects. It generally has low cytotoxicity against normal mammalian cells.

The antimicrobial peptide is generally highly cationic and hydrophobic. It typically contains several arginine and lysine residues, and it may not contain a single glutamate or aspartate. It usually contains a large proportion of hydrophobic residues. The peptide generally has an amphiphilic structure, with one surface being highly positive and the other hydrophobic.

The bioactive peptide and the encoding nucleotide sequence may be derived from plants, invertebrates, insects, amphibians and mammals, or from microorganisms such as bacteria and fungi.

The antimicrobial peptide may act on cell membranes of target microorganisms, e.g. through nonspecific binding to the membrane, usually in a membrane-parallel orientation, interacting only with one face of the bilayer.

The antimicrobial peptide typically has a structure belonging to one of five major classes: a helical, cystine-rich (defensin-like), b-sheet, peptides with an unusual composition of regular amino acids, and peptides containing uncommon modified amino acids.

Examples of alpha-helical peptides are Magainin 1 and 2; Cecropin A, B and P1; CAP18; Andropin; Clavanin A or AK; Styelin D and C; and Buforin II. Examples of cystine-rich peptides are a-Defensin HNP-1 (human neutrophil peptide) HNP-2 and HNP-3; b-Defensin-12, Drosomycin, g1-purothionin, and Insect defensin A. Examples of b-sheet peptides are Lactoferricin B, Tachyplesin I, and Protegrin PG1-5. Examples of peptides with an unusual composition are Indolicidin; PR-39; Bactenicin Bac5 and Bac7; and Histatin 5. Examples of peptides with unusual amino acids are Nisin, Gramicidin A, and Alamethicin.

Another example is the antifungal peptide (AFP) from Aspergillus giganteus. As explained in detail in WO 94/01459, which is hereby incorporated by reference, the antifungal polypeptide having the amino acid sequence shown in Fig. 1 has been found in several strains of the fungal species *A. giganteus*, an example of which is the *A. giganteus* strain deposited with the Centraallbureau voor Schimmelcultures (CBS) under the deposition number CBS 526.65.

However, the antifungal polypeptide, or variants thereof, suitable for the use according to the invention are expected to be derivable from other fungal species, especially other *Aspergillus* species such as *A. pallidus*, *A. clavatus*, *A. longivesica*, *A. rhizopodus* and *A. clavatonanicus*, because of the close relationship which exists between these species and *A. giganteus*.

In one embodiment of the invention the protein is an enzyme, such as glycosyl hydrolases, carbohydrases, peroxidases, proteases, lipases, phytases, polysaccharide lyases, oxidoreductases, transglutaminases and glycoseisomerases, in particular the following.

Parent Proteases

Parent proteases (i.e. enzymes classified under the Enzyme Classification number E.C. 3.4 in accordance with the Recommendations (1992) of the International Union of Biochemistry and Molecular Biology (IUBMB)) include proteases within this group.

Examples include proteases selected from those classified under the Enzyme Classification (E.C.) numbers:

3.4.11 (i.e. so-called aminopeptidases), including 3.4.11.5 (Prolyl aminopeptidase), 3.4.11.9 (X-pro aminopeptidase), 3.4.11.10 (Bacterial leucyl aminopeptidase), 3.4.11.12 (Thermophilic aminopeptidase), 3.4.11.15 (Lysyl aminopeptidase), 3.4.11.17 (Tryptophanyl aminopeptidase), 3.4.11.18 (Methionyl aminopeptidase).

3.4.21 (i.e. so-called serine endopeptidases), including 3.4.21.1 (Chymotrypsin), 3.4.21.4 (Trypsin), 3.4.21.25 (Cucumisin), 3.4.21.32 (Brachyurin), 3.4.21.48 (Cerevisin) and 3.4.21.62 (Subtilisin);

3.4.22 (i.e. so-called cysteine endopeptidases), including 3.4.22.2 (Papain), 3.4.22.3 (Ficain), 3.4.22.6 (Chymopapain), 3.4.22.7 (Asclepain), 3.4.22.14 (Actinidain), 3.4.22.30 (Caricain) and 3.4.22.31 (Ananain);

3.4.23 (i.e. so-called aspartic endopeptidases), including 3.4.23.1 (Pepsin A), 3.4.23.18 (Aspergillopepsin I), 3.4.23.20 (Penicillopepsin) and 3.4.23.25 (Saccharopepsin); and

3.4.24 (i.e. so-called metalloendopeptidases), including 3.4.24.28 (Bacillolysin).

Examples of relevant subtilisins comprise subtilisin BPN', subtilisin amylosacchariticus, subtilisin 168, subtilisin mesentericopeptidase, subtilisin Carlsberg, subtilisin DY, subtilisin 309, subtilisin 147, PD498 (WO 93/24623), thermitase, aqualysin, Bacillus PB92 protease, proteinase K, Protease TW7, and Protease TW3.

Specific examples of such readily available commercial proteases include Esperase®, Alcalase®, Neutrase®, Dyrazym®, Savinase® (WO 98/35026), Pyrase®, Pancreatic Trypsin NOVO (PTN), Bio-Feed™ Pro, Clear-Lens Pro (all enzymes available from Novo Nordisk A/S).

Examples of other commercial proteases include Maxtase®, Maxacal®, Maxapem® marketed by Gist-Brocades N.V., Opticlean® marketed by Solvay et Cie. and Purafect® marketed by Genencor International.

It is to be understood that also protease variants are contemplated as the parent protease. Examples of such protease variants are disclosed in EP 130.756 (Genentech), EP 214.435 (Henkel), WO 87/04461 (Amgen), WO 87/05050 (Genex), EP 251.446 (Genencor), EP 260.105 (Genencor), Thomas et al., (1985), Nature. 318, p. 375-376, Thomas et al., (1987), J. Mol. Biol., 193, pp. 803-813, Russel et al., (1987), Nature, 328, p. 496-500, WO 88/08028 (Genex), WO 88/08033 (Amgen), WO 89/06279 (Novo Nordisk A/S), WO 91/00345 (Novo Nordisk A/S), EP 525 610 (Solvay) and WO 94/02618 (Gist-Brocades N.V.).

The activity of proteases can be determined as described in "Methods of Enzymatic Analysis", third edition, 1984, Verlag Chemie, Weinheim, vol. 5.

Parent Lipases

Parent lipases (i.e. enzymes classified under the Enzyme Classification number E.C. 3.1.1 (Carboxylic Ester Hydrolases) in accordance with the Recommendations (1992) of the International Union of Biochemistry and Molecular Biology (IUBMB)) include lipases within this group.

Examples include lipases selected from those classified under the Enzyme Classification (E.C.) numbers:

3.1.1 (i.e. so-called Carboxylic Ester Hydrolases), including (3.1.1.3) Triacylglycerol lipases, (3.1.1.4.) Phosphorlipase A_2.

Examples of lipases include lipases derived from the following microorganisms. The indicated patent publications are incorporated herein by reference:

Humicola, e.g. H. brevispora, H. lanuginosa, H. brevis var. thermoidea and H. insolens (US 4,810,414).

Pseudomonas, e.g. Ps. fragi, Ps. stutzeri, Ps. cepacia and Ps. fluorescens (WO 89/04361), or Ps. plantarii or Ps. gladioli (US patent no. 4,950,417 (Solvay enzymes)) or Ps. alcaligenes and Ps. pseudoalcaligenes (EP 218 272) or Ps. mendocina (WO 88/09367; US 5,389,536).

Fusarium, e.g. F. oxysporum (EP 130,064) or F. solani pisi (WO 90/09446).

Mucor (also called Rhizomucor), e.g. M. miehei (EP 238023).

Chromobacterium (especially C. viscosum).

Aspergillus (especially A. niger).

Candida, e.g. C. cylindracea (also called C. rugosa) or C. antarctica (WO 88/02775) or C. antarctica lipase A or B (WO 94/01541 and WO 89/02916).

Geotricum, e.g. G. candidum (Schimada et al., (1989), J. Biochem., 106, 383-388).

Penicillium, e.g. P. camembertii (Yamaguchi et al., (1991), Gene 103, 61-67).

Rhizopus, e.g. R. delemar (Hass et al., (1991), Gene 109, 107-113) or R. niveus (Kugimiya et al., (1992) Biosci. Biotech. Biochem 56, 716-719) or R. oryzae.

Bacillus, e.g. B. subtilis (Dartois et al., (1993) Biochemica et Biophysica acta 1131, 253-260) or B. stearothermophilus (JP 64/7744992) or B. pumilus (WO 91/16422).

Specific examples of readily available commercial lipases include Lipolase® (WO 98/35026) Lipolase™ Ultra, Lipozyme®, Palatase®, Novozym® 435, Lecitase® (all available from Novo Nordisk A/S).

Examples of other lipases are Lumafast™, Ps. mendocian lipase from Genencor Int. Inc.; Lipomax™, Ps. pseudoalcaligenes lipase from Gist Brocades/Genencor Int. Inc.; Fusarium solani lipase (cutinase) from Unilever; Bacillus sp. lipase from Solvay enzymes. Other lipases are available from other companies.

It is to be understood that also lipase variants are contemplated as the parent enzyme. Examples of such are described in e.g. WO 93/01285 and WO 95/22615.

The activity of the lipase can be determined as described in "Methods of Enzymatic Analysis", Third Edition, 1984, Verlag Chemie, Weinhein, vol. 4, or as described in AF 95/5 GB (available on request from Novo Nordisk A/S).

Parent Oxidoreductases

Parent oxidoreductases (i.e. enzymes classified under the Enzyme Classification number E.C. 1 (Oxidoreductases) in accordance with the Recommendations (1992) of the International Union of Biochemistry and Molecular Biology (IUBMB)) include oxidoreductases within this group.

Examples include oxidoreductases selected from those classified under the Enzyme Classification (E.C.) numbers:
Glycerol-3-phosphate dehydrogenase _NAD+_ (1.1.1.8), Glycerol-3-phosphate dehydrogenase _NAD(P)$^+$_ (1.1.1.94), Glycerol-3-phosphate 1-dehydrogenase _NADP_ (1.1.1.94), Glucose oxidase (1.1.3.4), Hexose oxidase (1.1.3.5), Catechol oxidase (1.1.3.14), Bilirubin oxidase (1.3.3.5), Alanine dehydrogenase (1.4.1.1), Glutamate dehydrogenase (1.4.1.2), Glutamate dehydrogenase _NAD(P)$^+$_ (1.4.1.3), Glutamate dehydrogenase _NADP$^+$_ (1.4.1.4), L-Amino acid dehydrogenase (1.4.1.5), Serine dehydrogenase (1.4.1.7), Valine dehydrogenase _NADP$^+$_ (1.4.1.8), Leucine dehydrogenase (1.4.1.9), Glycine dehydrogenase (1.4.1.10), L-Amino-acid oxidase (1.4.3.2.), D-Amino-acid oxidase(1.4.3.3), L-Glutamate oxidase (1.4.3.11), Protein-lysine 6-oxidase (1.4.3.13), L-lysine oxidase (1.4.3.14), L-Aspartate oxidase (1.4.3.16), D-amino-acid dehydrogenase (1.4.99.1), Protein disulfide reductase (1.6.4.4), Thioredoxin reductase (1.6.4.5), Protein disulfide reductase (glutathione) (1.8.4.2), Laccase (1.10.3.2) (SEQ ID NO. 6), Catalase (1.11.1.6), Peroxidase

(1.11.1.7), Lipoxygenase (1.13.11.12), Superoxide dismutase (1.15.1.1)

Said Glucose oxidases may be derived from Aspergillus niger.

Said Laccases may be derived from Polyporus pinsitus, Myceliophtora thermophila, Coprinus cinereus, Rhizoctonia solani, Rhizoctonia praticola, Scytalidium thermophilum and Rhus vernicifera.

Bilirubin oxidases may be derived from Myrothechecium verrucaria.

The Peroxidase may be derived from e.g. Soy bean, Horseradish or Coprinus cinereus.

The Protein Disulfide reductase may be any of the mentioned in DK patent applications No. 768/93, 265/94 and 264/94 (Novo Nordisk A/S), which are hereby incorporated as references, including Protein Disulfide reductases of bovine origin, Protein Disulfide reductases derived from Aspergillus oryzae or Aspergillus niger, and DsbA or DsbC derived from Escherichia coli.

Specific examples of readily available commercial oxidoreductases include Gluzyme™ (enzyme available from Novo Nordisk A/S). However, other oxidoreductases are available from others. It is to be understood that also variants of oxidoreductases are contemplated as the parent enzyme.

The activity of oxidoreductases can be determined as described in "Methods of Enzymatic Analysis", third edition, 1984, Verlag Chemie, Weinheim, vol. 3.

Parent Carbohydrases

Parent carbohydrases may be defined as all enzymes capable of breaking down carbohydrate chains (e.g. starches) of especially five and six member ring structures (i.e. enzymes classified under the Enzyme Classification number E.C. 3.2 (glycosidases) in accordance with the Recommendations (1992) of the International Union of Biochemistry and Molecular Biology (IUBMB)). Also included in the group of carbohydrases according to the invention are enzymes capable of isomerizing carbohydrates e.g. six member ring structures, such as D-glucose to e.g. five member ring structures like D-fructose.

Examples include carbohydrases selected from those classified under the Enzyme Classification (E.C.) numbers:

α-amylase (3.2.1.1) β-amylase (3.2.1.2), glucan 1,4-α-glucosidase (3.2.1.3), cellulase (3.2.1.4), endo-1,3(4)-β-glucanase (3.2.1.6), endo-1,4-β-xylanase (3.2.1.8), dextranase (3.2.1.11), chitinase (3.2.1.14), polygalacturonase (3.2.1.15), lysozyme (3.2.1.17), β-glucosidase (3.2.1.21), α-galactosidase (3.2.1.22), β-galactosidase (3.2.1.23), amylo-1,6-glucosidase (3.2.1.33), xylan 1,4-β-xylosidase (3.2.1.37), glucan endo-1,3-β-D-glucosidase (3.2.1.39), α-dextrin endo-1,6-glucosidase (3.2.1.41), sucrose α-glucosidase (3.2.1.48), glucan endo-1,3-α-glucosidase (3.2.1.59), glucan 1,4-β-glucosidase (3.2.1.74), glucan endo-1,6-β-glucosidase (3.2.1.75), arabinan endo-1,5-α-arabinosidase (3.2.1.99), lactase (3.2.1.108), chitonanase (3.2.1.132) and xylose isomerase (5.3.1.5).

Examples of relevant carbohydrases include α-1,3-glucanases derived from Trichoderma harzianum; α-1,6-glucanases derived from a strain of Paecilomyces; β-glucanases derived from Bacillus subtilis; β-glucanases derived from Humicola insolens; β-glucanases derived from Aspergillus niger; β-glucanases derived from a strain of Trichoderma; β-glucanases derived from a strain of Oerskovia xanthineolytica; exo-1,4-α-D-glucosidases (glucoamylases) derived from Aspergillus niger; α-amylases derived from Bacillus subtilis; α-amylases derived from Bacillus amyloliquefaciens; α-amylases derived from Bacillus stearothermophilus; α-amylases derived from Aspergillus oryzae; α-amylases derived from non-pathogenic microorganisms; α-galactosidases derived from Aspergillus niger; Pentosanases, xylanases, cellobiases, cellulases, hemi-cellulases deriver from Humicola insolens; cellulases derived from Trichoderma reesei; cellulases derived from non-pathogenic mold; pectinases, cellulases, arabinases, hemicelluloses derived from Aspergillus niger; dextranases derived from Penicillium lilacinum; endo-glucanase derived from non-pathogenic mold; pullulanases derived from Bacillus acidopullyticus; β-galactosidases derived from Kluyveromyces fragilis; xylanases derived from Trichoderma reesei;

Specific examples of readily available commercial carbohydrases include Alpha-Gal™, Bio-Feed™ Alpha, Bio-Feed™ Beta, Bio-Feed™ Plus, Bio-Feed™ Plus, Novozyme® 188, Carezyme® (SEQ ID NO. 5), Celluclast®, Cellusoft®, Ceremyl®, Citrozym™, Denimax™, Dezyme™, Dextrozyme™, Finizym®, Fungamyl™, Gamanase™, Glucanex®, Lactozym®, Maltogenase™, Pentopan™, Pectinex™, Promozyme®, Pulpzyme™, Novamyl™, Termamyl®, AMG (Amyloglucosidase

Novo), Maltogenase®, Sweetzyme®, Aquazym®, Natalase® (SEQ ID NO. 4) (all enzymes available from Novo Nordisk A/S). Other carbohydrases are available from other companies.

It is to be understood that also carbohydrase variants are contemplated as the parent enzyme.

The activity of carbohydrases can be determined as described in "Methods of Enzymatic Analysis", third edition, 1984, Verlag Chemie, Weinheim, vol. 4.

Parent Transferases

Parent transferases (i.e. enzymes classified under the Enzyme Classification number E.C. 2 in accordance with the Recommendations (1992) of the International Union of Biochemistry and Molecular Biology (IUBMB)) include transferases within this group.

The parent transferases may be any transferase in the subgroups of transferases: transferases transferring one-carbon groups (E.C. 2.1); transferases transferring aldehyde or residues (E.C 2.2); acyltransferases (E.C. 2.3); glucosyltransferases (E.C. 2.4); transferases transferring alkyl or aryl groups, other that methyl groups (E.C. 2.5); transferases transferring nitrogeneous groups (2.6).

In a preferred embodiment the parent transferase is a transglutaminase E.C 2.3.2.13(Protein-glutamine μ-glutamyltransferase).

Transglutaminases are enzymes capable of catalyzing an acyl transfer reaction in which a gamma-carboxyamide group of a pep-

tide-bound glutamine residue is the acyl donor. Primary amino groups in a variety of compounds may function as acyl acceptors with the subsequent formation of monosubstituted gamma-amides of peptide-bound glutamic acid. When the epsilon-amino group of a lysine residue in a peptide-chain serves as the acyl acceptor, the transferases form intramolecular or intermolecular gamma-glutamyl-epsilon-lysyl crosslinks.

Examples of transglutaminases are described in the pending DK patent application no. 990/94 (Novo Nordisk A/S).

The parent transglutaminase may be of human, animal (e.g. bovine) or microbial origin.
Examples of such parent transglutaminases are animal derived Transglutaminase, FXIIIa; microbial transglutaminases derived from Physarum polycephalum (Klein et al., Journal of Bacteriology, Vol. 174, p. 2599-2605); transglutaminases derived from Streptomyces sp., including Streptomyces lavendulae, Streptomyces lydicus (former Streptomyces libani) and Streptoverticillium sp., including Streptoverticillium mobaraense, Streptoverticillium cinnamoneum, and Streptoverticillium griseocarneum (Motoki et al., US 5,156,956; Andou et al., US 5,252,469; Kaempfer et al., Journal of General Microbiology, Vol. 137, p. 1831-1892; Ochi et al., International Journal of Sytematic Bacteriology, Vol. 44, p. 285-292; Andou et al., US 5,252,469; Williams et al., Journal of General Microbiology, Vol. 129, p. 1743-1813).

It is to be understood that also transferase variants are contemplated as the parent enzyme.

The activity of transglutaminases can be determined as described in "Methods of Enzymatic Analysis", third edition, 1984, Verlag Chemie, Weinheim, vol. 1-10.

Parent Phytases

Parent phytases are included in the group of enzymes classified under the Enzyme Classification number E.C. 3.1.3 (Phosphoric Monoester Hydrolases) in accordance with the Recommendations (1992) of the International Union of Biochemistry and Molecular Biology (IUBMB)).

Phytases are enzymes produced by microorganisms which catalyse the conversion of phytate to inositol and inorganic phosphorus

Phytase producing microorganisms comprise bacteria such as Bacillus subtilis, Bacillus natto and Pseudomonas; yeasts such as Saccharomyces cerevisiae; and fungi such as Aspergillus niger, Aspergillus ficuum, Aspergillus awamori, Aspergillus oryzae, Aspergillus terreus or Aspergillus nidulans, and various other Aspergillus species).

Examples of parent phytases include phytases selected from those classified under the Enzyme Classification (E.C.) numbers: 3-phytase (3.1.3.8) and 6-phytase (3.1.3.26).

The activity of phytases can be determined as described in "Methods of Enzymatic Analysis", third edition, 1984, Verlag Chemie, Weinheim, vol. 1-10, or may be measured according to the method described in EP-A1-0 420 358, Example 2 A.

Lyases

Suitable lyases include Polysaccharide lyases: Pectate lyases (4.2.2.2) and pectin lyases (4.2.2.10), such as those from Bacillus licheniformis disclosed in WO 99/27083.

Isomerases

Protein Disulfide Isomerase.
Without being limited thereto suitable protein disulfide isomerases include PDIs described in WO 95/01425 (Novo Nordisk A/S) and suitable glucose isomerases include those described in Biotechnology Letter, Vol. 20, No 6, June 1998, pp. 553-56.
Contemplated isomerases include xylose/glucose Isomerase (5.3.1.5) including Sweetzyme®.

The environmental allergens that are of interest for epitope mapping include allergens from pollen, dust mites, mammals, venoms, fungi, food items, and other plants.

Pollen, allergens include but are not limited to those of the order Fagales, Oleales, Pinales, Poales, Asterales, and Urticales; including those from Betula, Alnus, Corylus, Carpinus, Olea, Phleum pratense and Artemisia vulgaris, such as Aln g1, Cor a1, Car b1, Cry j1, Amb a1 and a2, Art v1, Par j1, Ole e1, Ave v1, and Bet v1 (WO 99/47680).

Mite allergens include but are not limited to those from Derm. farinae and Derm. pteronys., such as Der f1 and f2, and Der p1 and p2.

From mammals, relevant environmental allergens include but are not limited to those from cat, dog, and horse as well as from dandruff from the hair of those animals, such as Fel d1, Can f1 Equ c1, c2, c3.

Venum allergens include but are not limited to PLA2 from bee venom as well as Apis m1 and m2, Ves g1, g2 and g5, and te Pol and Sol allergens.

Fungal allergens include those from Alternaria alt. and Cladospo. herb. such as Alt a1 and Cla h1.

Food allergens include but are not limited to those from milk (lactoglobulin), egg (ovalbumin), peanuts, hazelnuts, wheat (alfa-amylase inhibitor),
Other plant allergens include latex (hevea brasiliensis).

In addition, a number of proteins of interest for expression in transgenic plants could be useful objects for epitope engineering. If for instance a heterologous enzyme is introduced into a transgenic plant e.g. to increase the nutritional value of food or feed derived from that plant, that enzyme may lead to allergenicity problems in humans or animals ingesting the plant-derived material. Epitope mapping and engineering of such heterologous enzymes or other proteins of transgenic plants may lead to reduction or elimination of this problem. Hence, the methods of this patent are also useful for potentially modifying proteins for heterologous expression in plants and plant cells.

Materials and methods

Materials

ELISA reagents:

Horse Radish Peroxidase labelled pig anti-rabbit-Ig (Dako, DK, P217, dilution 1:1000).

Rat anti-mouse IgE (Serotec MCA419; dilution 1:100).

Mouse anti-rat IgE (Serotec MCA193; dilution 1:200).

Biotin-labelled mouse anti-rat IgG1 monoclonal antibody (Zymed 03-9140; dilution 1:1000)

Biotin-labelled rat anti-mouse IgG1 monoclonal antibody (Serotec MCA336B; dilution 1:2000)

Streptavidin-horse radish peroxidase (Kirkegård & Perry 14-30-00; dilution 1:1000).

Buffers and Solutions:

- PBS (pH 7.2 (1 liter))

NaCl	8.00 g
KCl	0.20 g
K_2HPO_4	1.04 g
KH_2PO_4	0.32 g

- Washing buffer PBS, 0.05% (v/v) Tween 20
- Blocking buffer PBS, 2% (wt/v) Skim Milk powder
- Dilution buffer PBS, 0.05% (v/v) Tween 20, 0.5% (wt/v) Skim Milk powder
- Citrate buffer 0.1M, pH 5.0-5.2
- Stop-solution (DMG-buffer)
- Sodium Borate, borax (Sigma)
- 3,3-Dimethyl glutaric acid (Sigma)
- Tween 20: Poly oxyethylene sorbitan mono laurate (Merck cat no. 822184)

- PMSF (phenyl methyl sulfonyl flouride) from Sigma
- Succinyl-Alanine-Alanine-Proline-Phenylalanine-paranitro-anilide (Suc-AAPF-pNP) Sigma no. S-7388, Mw 624.6 g/mol.
- mPEG (Fluka)

Colouring substrate:

OPD: o-phenylene-diamine, (Kementec cat no. 4260)

Methods

Immunisation of Brown Norway rats:

Twenty intratracheal (IT) immunisations were performed weekly with 0,100 ml 0.9% (wt/vol) NaCl (control group), or 0,100 ml of a protein dilution (~0,1-1 mg/ml). Each group contained 10 rats. Blood samples (2 ml) were collected from the eye one week after every second immunisation. Serum was obtained by blood clothing and centrifugation and analysed as indicated below.

Immunisation of Balb/C mice:

Twenty subcutaneous (SC) immunisations were performed weekly with 0.05 ml 0.9% (wt/vol) NaCl (control group), or 0,050 ml of a protein dilution (~0,01-0,1 mg/ml). Each group contained 10 female Balb/C mice (about 20 grams) purchased from Bomholdtgaard, Ry, Denmark. Blood samples (0,100 ml) were collected from the eye one week after every second immunisation. Serum was obtained by blood clothing and centrifugation and analysed as indicated below.

ELISA Procedure for detecting serum levels of IgE and IgG:

Specific IgG1 and IgE levels were determined using the ELISA specific for mouse or rat IgG1 or IgE. Differences between data sets were analysed by using appropriate statistical methods.

Activation of CovaLink plates:

A fresh stock solution of cyanuric chloride in acetone (10 mg/ml) is diluted into PBS, while stirring, to a final concentration of 1 mg/ml and immediately aliquoted into CovaLink NH2 plates (100 microliter per well) and incubated for 5 minutes at room temperature. After three washes with PBS, the plates are dryed at 50°C for 30 minutes, sealed with sealing tape, and stored in plastic bags at room temperature for up to 3 weeks.

Mouse anti-Rat IgE was diluted 200x in PBS (5 microgram/ml). 100 microliter was added to each well. The plates were coated overnight at 4 °C.

Unspecific adsorption was blocked by incubating each well for 1 hour at room temperature with 200 microliter blocking buffer. The plates were washed 3x with 300 microliter washing buffer.

Unknown rat sera and a known rat IgE solution were diluted in dilution buffer: Typically 10x, 20x and 40x for the unknown sera, and ½ dilutions for the standard IgE starting from 1 µg/ml. 100 microliter was added to each well. Incubation was for 1 hour at room temperature.

Unbound material was removed by washing 3x with washing buffer. The anti-rat IgE (biotin) was diluted 2000x in dilution buffer. 100 microliter was added to each well. Incubation was for 1 hour at room temperature. Unbound material was removed by washing 3x with washing buffer.

Streptavidin was diluted 1000x in dilution buffer. 100 microliter was added to each well. Incubation was for 1 hour at room temperature. Unbound material was removed by washing 3x with 300 microliter washing buffer. OPD (0.6 mg/ml) and H_2O_2 (0.4 microliter /ml) were dissolved in citrate buffer. 100 microliter was added to each well. Incubation was for 30 minutes at room temperature. The reaction was stopped by addition of 100 microliter H_2SO_4. The plates were read at 492 nm with 620 nm as reference.

Similar determination of IgG can be performed using anti Rat-IgG and standard rat IgG reagents.

Similar determinations of IgG and IgE in mouse serum can be performed using the corresponding species-specific reagents.

Direct IgE assay:

To determine the IgE binding capacity of protein variants one can use an assay, essentially as described above, but using sequential addition of the follwing reagents:

1) Mouse anti-rat IgE antibodies coated in wells;
2) Known amounts of rat antiserum containing igE against the parent protein;

3) Dilution series of the protein variant in question (or parent protein as positive control);
4) Rabbit anti-parent antibodies
5) HRPO-labelled anti-rabbit Ig antibodies for detection using OPD as described.

The relative IgE binding capacity (end-point and/or affinity) of the protein variants relative to that of the parent protein are determined from the dilution-response curves. The IgE-positive serum can be of other animals (including humans that inadvertently have been senstitized to the parent protein) provided that the species-specific anti-IgE capture antibodies are changed accordingly.

Competitive ELISA (C-ELISA):

C-ELISA was performed according to established procedures. In short, a 96 well ELISA plate was coated with the parent protein. After proper blocking and washing, the coated antigen was incubated with rabbit anti-enzyme polyclonal antiserum in the presence of various amounts of modified protein (the competitior). The residual amount of rabbit antiserum was detected by horseraddish peroxidase-labelled pig anti-rabbit immunoglobulin.

Protein sequences and alignments:

PD498: The sequence of PD498 is provided in Table 1. The numeration used throughout this invention is the BPN' notation based on the alignment of Table 1. The protein structure of PD498 is disclosed in WO98/35026 (Novo Nordisk).

The sequence of Savinase is provided in Table 1. The numeration used throughout this invention is the BPN' notation based on the alignment of Table 1. The protein structure of PD498 is disclosed in WO98/35026 (Novo Nordisk). (von der Osten et al., (1993), Journal of Biotechnology, 28, p. 55), and the structure can be found in BETZEL et al, J.MOL.BIOL., Vol. 223, p. 427, 1992 (1svn.pdb).

The subtilisin sequences are aligned to the BPN' sequence of Subtilisin Novo from B.amyloliquefaciens as indicated in Table 1. All amino acid residue numbers for Savinase, PD498 and other subtilisins are given in the BPN' numeration and can be converted back using Table 1 below.

The lipolase sequence is provided in SEQ ID NO 1 and the structure is disclosed in WO 98/35026.

Amylase:

The amylase used in the examples is the alpha-amylase of Bacillus halmapalus (WO96/23873), which is called amylase SP722 (the wild-type). Its sequence is shown in SEQ ID NO 2 and the corresponding protein structure was built from the BA2 structure, as described in WO96/23874. The first four amino acids of the structural model are not defined, hence the sequence used for numeration of amino acid residues in the examples of this invention is four amino acids shorter than the one of the full length protein SP722.

Several variants of this amylase are available (WO96/23873). One particularly useful variant has deleted two amino acid residues

at D-G at positions 183 and 184 of the SEQ ID NO 2 (corresponding to residues 179 and 180 of the modelled structure). This variant is called JE-1 or Natalase.

Another amylase that is particularly useful is the amylase AA560: This alkaline a-amylase may be derived from a strain of Bacillus sp. DSM 12649. The strain was deposited on 25th January 1999 by the assignee under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure at Deutshe Sammmlung von Microorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1b, D-38124 Braunschweig DE.

Laccase:

The laccase used in this invention is that from Myceliophthora thermophila (WO98/38287), the sequence of which is shown as SEQ ID NO 3. The structure was built by homology modeling to the Coprinus cinereus laccase as shown in WO98/38287.

Cellulase:

The cellulase sequence and structure used in the present invention is that of the core fragment of endoglucanase V from Humicola insolens (aka Cel45 or Carezyme). The core fragment structure is available as 3eng.pdb (G.J.DAVIES et al. ACTA CRYSTALLOGR.,SECT.D, Vol. 52, p.7 1996; G.J.DAVIES et al. BIOCHEMISTRY, V. 34, p. 16210, 1995); SwissProt accession number P43316, and the sequences shown in SEQ ID 4. The corresponding full-length

sequence is disclosed in WO91/17243 and shown here in SEQ ID NO 5. The numeration of all description and claims of this invention pertain to the core fragment, however, it is contemplated that all claims are also valid for the corresponding positions in the full-length protein.

Table 1: Alignment and numeration scheme for subtilisins.

BPN'	Savinase	PD498	BPN'	Savinase	PD498	BPN'	Savinase	PD498
-6		1 W	85 A	83 A	94 A	181 D	175 D	186 D
-5		2 S	86 P	84 P	95 P	182 S	176 Q	187 S
-4		3 P	87 S	85 S	96 D	183 S	177 N	188 N
-3		4 N	88 A	86 A	97 T	184 N	178 N	189 D
-2		5 D	89 S	87 E	98 K	185 Q	179 N	190 R
-1		6 P	90 L	88 L	99 I	186 R	180 R	191 K
1 A	1 A	7 Y	91 Y	89 Y	100 L	187 A	181 A	192 A
2 Q	2 Q	8 Y	92 A	90 A	101 A	188 S	182 S	193 S
3 S	3 S	9 S	93 V	91 V	102 V	189 F	183 F	194 F
*3a		10 A	94 K	92 K	103 R	190 S	184 S	195 S
4 V	4 V	11 Y	95 V	93 V	104 V	191 S	185 Q	196 N
5 P	5 P	12 Q	96 L	94 L	105 L	192 V	186 Y	197 Y
6 Y	6 W	13 Y	97 G	95 G	106 D	193 G	187 G	198 G
7 G	7 G	14 G	98 A	96 A	107 A	194 P	188 A	199 T
8 V	8 I	15 P	99 D	97 S	108 N	195 E	189 G	200 W
9 S	9 S	16 Q	100 G	98 G	109 G	196 L	190 L	201 V
10 Q	10 R	17 N	101 S	99 S	110 S	197 D	191 D	202 D
11 I	11 V	18 T	102 G	100 G	111 G	198 V	192 I	203 V
12 K	12 Q	19 S	103 Q	101 S	112 S	199 M	193 V	204 T
13 A	13 A	20 T	104 Y	102 V	113 L	200 A	194 A	205 A
14 P	14 P	21 P	105 S	103 S	114 D	201 P	195 P	206 P
15 A	15 A	22 A	106 W	104 S	115 S	202 G	196 G	207 G
16 L	16 A	23 A	107 I	105 I	116 I	203 V	197 V	208 V
17 H	17 H	24 W	108 I	106 A	117 A	204 S	198 N	209 N
18 S	18 N	25 D	109 N	107 Q	118 S	205 I	199 V	210 I
19 Q	19 R	26 V	110 G	108 G	119 G	206 Q	200 Q	211 A
20 G	20 G		111 I	109 L	120 I	207 S	201 S	212 S
21 Y	21 L	27 T	112 E	110 E	121 R	208 T	202 T	213 T
22 T	22 T	28 R	113 W	111 W	122 Y	209 L	203 Y	214 V
23 G	23 G	29 G	114 A	112 A	123 A	210 P	204 P	215 P
24 S	24 S	30 S	115 I	113 G	124 A	211 G	205 G	216 N
25 N	25 G	31 S	116 A	114 N	125 D	212 N	206 S	217 N
26 V	26 V	32 T	117 N	115 N	126 Q	213 K	207 T	218 G
27 K	27 K	33 Q	118 N	116 G	127 G	214 Y	208 Y	219 Y
28 V	28 V	34 T	119 M	117 M	128 A	215 G	209 A	220 S
*28a		35 V	120 D	118 H	129 K	216 A	210 S	221 Y
29 A	29 A	36 A	121 V	119 V	130 V	217 Y	211 L	222 M
30 V	30 V	37 V	122 I	120 A	131 L	218 N	212 N	223 S
31 I	31 L	38 L	123 N	121 N	132 N	219 G	213 G	224 G
32 D	32 D	39 D	124 M	122 L	133 L	220 T	214 T	225 T
33 S	33 T	40 S	125 S	123 S	134 S	221 S	215 S	226 S
34 G	34 G	41 G	126 L	124 L	135 L	222 M	216 M	227 M
35 I	35 I	42 V	127 G	125 G	136 G	223 A	217 A	228 A
36 D	*	43 D	128 G	126 S	137 C	224 S	218 T	229 S
37 S	36 S	44 Y	129 P	127 P	138 E	225 P	219 P	230 P
38 S	37 T	45 N	130 S	128 S	139 C	226 H	220 H	231 H
39 H	38 H	46 H	131 G	129 P	140 N	227 V	221 V	232 V
40 P	39 P	47 P	132 S	130 S	141 S	228 A	222 A	233 A
41 D	40 D	48 D	133 A	131 A	142 T	229 G	223 G	234 G
42 L	41 L	49 L	134 A	132 T	143 T	230 A	224 A	235 L
43 K	42 N	50 A	135 L	133 L	144 L	231 A	225 A	236 A
44 V	43 I	51 R	136 K	134 E	145 K	232 A	226 A	237 A
*44a		52 K	137 A	135 Q	146 S	233 L	227 L	238 L

*44aa		53 V	138 A	136 A	147 A	234 I	228 V	239 L
45 A	44 R	54 I	139 V	137 V	148 V	235 L	229 K	240 A
46 G	45 G	55 K	140 D	138 N	149 D	236 S	230 Q	241 S
47 G	46 G	56 G	141 K	139 S	150 Y	237 K	231 K	242 Q
48 A	47 A	57 Y	142 A	140 A	151 A	238 H	232 N	243 G
49 S	48 S	58 D	143 V	141 T	152 W	239 P	233 P	244 K
50 M	49 F	59 F	144 A	142 S	153 N	240 N	234 S	245 N
51 V	50 V	60 I	145 S	143 R	154 K	241 W	235 W	
52 P	51 P	61 D	146 G	144 G	155 G	242 T	236 S	
53 S	52 G	62 R	147 V	145 V	156 A	243 N	237 N	246 N
54 E	53 E	63 D	148 V	146 L	157 V	244 T	238 V	247 V
55 T	54 P	64 N	149 V	147 V	158 V	245 Q	239 Q	248 Q
56 N	*	65 N	150 V	148 V	159 V	246 V	240 I	249 I
57 P	55 S	66 P	151 A	149 A	160 A	247 R	241 R	250 R
58 F	56 T	67 M	152 A	150 A	161 A	248 S	242 N	251 Q
59 Q	57 Q		153 A	151 S	162 A	249 S	243 H	252 A
60 D	58 D	68 D	154 G	152 G	163 G	250 L	244 L	253 I
61 N	59 G	69 L	155 N	153 N	164 N	251 E	245 K	254 E
62 N	60 N	70 N	156 E	154 S	165 D	252 N	246 N	255 Q
63 S	61 G	71 G	157 G	155 G	166 N	253 T	247 T	256 T
64 H	62 H	72 H	158 T	156 A	167 V	254 T	248 A	257 A
65 G	63 G	73 G	159 S	*		255 T	249 T	258 D
66 T	64 T	74 T	160 G	157 G	168 S	256 K	250 S	259 K
67 H	65 H	75 H	161 S	158 S	169 R	257 L	251 L	260 I
68 V	66 V	76 V	162 S	159 I	170 T	258 G	252 G	261 S
69 A	67 A	77 A	163 S	160 S	171 F	259 D	253 S	262 G
70 G	68 G	78 G	164 T	*		260 S	254 T	263 T
71 T	69 T	79 T	165 V	*		261 F	255 N	264 G
72 V	70 I	80 V	166 G	*		262 Y	256 L	265 T
73 A	71 A	81 A	167 Y	161 Y	172 Q	263 Y	257 Y	266 N
74 A	72 A	82 A	168 P	162 P	173 P	264 G	258 G	267 F
75 L	73 L	83 D	169 G	163 A	174 A	*264a		268 K
*75a		84 T	170 K	164 R	175 S	265 K	259 S	269 Y
76 N	74 N	85 N	171 Y	165 Y	176 Y	266 G	260 G	270 G
77 N	75 N	86 N	172 P	166 A	177 P	267 L	261 L	271 K
78 S	76 S	87 G	173 S	167 N	178 N	268 I	262 V	272 I
79 I	77 I	88 I	174 V	168 A	179 A	269 N	263 N	273 N
80 G	78 G	89 G	175 I	169 M	180 I	270 V	264 A	274 S
81 V	79 V	90 V	176 A	170 A	181 A	271 Q	265 E	275 N
82 L	80 L	91 A	177 V	171 V	182 V	272 A	266 A	276 K
83 G	81 G	92 G	178 G	172 G	183 G	273 A	267 A	277 A
84 V	82 V	93 M	179 A	173 A	184 A	274 A	268 T	278 V
			180 V	174 T	185 I	275 Q	269 R	279 R
								280 Y

Examples

Example 1

Identification of epitope sequences and epitope patterns.

High diversity libraries (10^{12}) of phages expressing random hexa-, nona- or dodecapetides as part of their membrane proteins, were screened for their capacity to bind purified spe-

cific rabbit IgG, and purified rat and mouse IgG1 and IgE antibodies. The phage libraries were obtained according to prior art (se WO 9215679 hereby incorporated by reference).

The antibodies were raised in the respective animals by subcutaneous, intradermal, or intratracheal injection of relevant proteins (e.g. proteases, lipases, amylases, oxidoreductases) dissolved in phosphate buffered saline (PBS). The respective antibodies were purified from the serum of immunised animals by affinity chromatography using paramagnetic immunobeads (Dynal AS) loaded with pig anti-rabbit IgG, mouse anti-rat IgG1 or IgE, or rat anti-mouse IgG1 or IgE antibodies.

The respective phage libraries were incubated with the IgG, IgG1 and IgE antibody coated beads. Phages, which express oligopeptides with affinity for rabbit IgG, or rat or mouse IgG1 or IgE antibodies, were collected by exposing these paramagnetic beads to a magnetic field. The collected phages were eluted from the immobilised antibodies by mild acid treatment, or by elution with intact enzyme. The isolated phages were amplified as know to the specialist. Alternatively, immobilised phages were directly incubated with E.coli for infection. In short, F-factor positive E.coli (e.g. XL-1 Blue, JM101, TG1) were infected with M13-derived vector in the presence of a helper-phage (e.g. M13K07), and incubated, typically in 2xYT containing glucose or IPTG, and appropriate antibiotics for selection. Finally, cells were removed by centrifugation. This cycle of events was repeated 2-5 times on the respective cell supernatants. After selection round 2, 3, 4, and 5, a fraction of the infected E.coli was incubated on selective 2xYT agar plates, and the specificity of the emerging phages

was assessed immunologically. Thus, phages were transferred to a nitrocellulase (NC) membrane. For each plate, 2 NC-replicas were made. One replica was incubated with the selection antibodies, the other replica was incubated with the selection antibodies and the immunogen used to obtain the antibodies as competitor. Those plaques that were absent in the presence of immunogen, were considered specific, and were amplified according to the procedure described above.

The specific phage-clones were isolated from the cell supernatant by centrifugation in the presence of polyethylenglycol. DNA was isolated, the DNA sequence coding for the oligopeptide was amplified by PCR, and the DNA sequence was determined, all according to standard procedures. The amino acid sequence of the corresponding oligopeptide was deduced from the DNA sequence.

Thus, a number of peptide sequences with specificity for the protein specific antibodies, described above, were obtained. These sequences were collected in a database, and analysed by sequence alignment to identify epitope patterns. For this sequence alignment, conservative substitutions (e.g. aspartate for glutamate, lysine for arginine, serine for threonine) were considered as one. This showed that most sequences were specific for the protein the antibodies were raised against. However, several cross-reacting sequences were obtained from phages that went through 2 selection rounds only. In the first round 22 epitope patterns were identified.

In further rounds of phage display, more antibody binding sequences were obtained leading to more epitope patterns. Further, the literature was searched for peptide sequences that

have been found to bind environmental allergen-specific antibodies (J All Clin Immunol 93 (1994) pp. 34-43; Int Arch Appl Immunol 103 (1994) pp. 357-364; Clin Exp Allergy 24 (1994) pp. 250-256; Mol Immunol 29 (1992) pp. 1383-1389; J Immunol 121 (1989) pp. 275-280; J. Immunol 147 (1991) pp. 205-211; Mol Immunol 29 (1992) pp. 739-749; Mol Immunol 30 (1993) pp. 1511-1518; Mol Immunol 28 (1991) pp. 1225-1232; J. Immunol 151 (1993) pp. 7206-7213). These antibody binding peptide sequences were included in the database.

The current database of antibody binding peptides identified and their corresponding epitope patterns are shown in Table 2-7 below.

Tables 2-7: Overview of the antibody binding peptide sequences, epitope patterns and epitope sequences. The type of antibody used for identifying the antibody binding sequences is indicated as IgG or IgE and the species from which the antibodies were derived are indicated as mo (mouse), ra (rat) and hu (human).

Table 2: Savinase antibody binding peptide sequences, epitope patterns and epitope sequences.

VQVYGDTSA	Phage display	Q > Y > D >	savinase	savinase	Q206 V81 Y214 G80 D41 T208	sav1.1	Ra	
LQCVGS	Protein fragments		a-amylase inhibitor	savinase	L21 Q236 V26 G25 S24	sav19.1		Hu
KRFANTELA	Phage display	R/K R F > N	savinase	savinase	K251 R247 A174 N173	sav6.1	Ra-Mo	Mo
LDQIFFTRW	Phage display	D/E Q I F F T	savinase	savinase	L42/L75 D41 Q2 I79	sav5.1	Ra	
FNDAFFVKM	Phage display		savinase	savinase	N185 D181 A187 F189 V203	sav11.0	Ra	
ANIPIWSRSA	Phage display	> R S A	savinase	savinase	R145 S144 A142	sav3.2-lac1.0-lip4.0-pd5.0	Ra	
ANIPIWSRSA	Phage display	> R S A	savinase	savinase	S188 R186 S190 A179	sav3.1-lac1.0-lip4.0-pd5.0	Ra	
RQSTDFGTT	Phage display	R Q > > D/E	savinase	savinase	R186 Q191 S156	sav2.2	Ra	
VQVYGDTSA	Phage display	Q > Y > D >	savinase	savinase	Q191 Y192 G193/A194/G195 D197 S265	sav1.2	Ra	
RRFSNATRA	Phage display	R/K R F > N	savinase	savinase	K251 R247 A174 N173	sav6.1	Ra-Mo	Mo
CTARLRAGNACG	Phage display	A R > A	savinase	savinase	A172/A169 R170 A194 G193 N261	sav10.4	Ra	
LDQIFFTRW	Phage display	D/E Q I F F T	savinase	savinase	D60 Q59 I44/I35	sav5.2	Ra	
LDQIFFTRW	Phage display	D/E Q I F F T	savinase	savinase	L42/L75 D41 Q2 I79	sav5.1	Ra	
EQIFFTSGL	Phage display	D/E Q I F F T	savinase	savinase	E112 Q109 I79	sav5.4	Ra	
GRFSNSKFK	Phage display	L > G R S	savinase	savinase	L196 G195 R170 S163	sav9.2-je4.0-lip5.1-5.2	Ra	
AVLRDC	Protein fragments		a-amylase inhibitor	savinase	A254 V268 L267 R10 D181	sav18.1-pd18.1-18.2		Hu
LQCVGS	Protein fragments		a-amylase inhibitor	savinase	L217 Q206 V81 G80 S3	sav19.2		Hu
LRQCNERCV	Phage display	R Q > > D/E	savinase	savinase	L267 R10 Q12 N269 E271 R275	sav2.1	Ra	
SPVTKRASLKIDSKK	Protein fragments		Der p II	savinase	A88 S87/T22 L233 K235 I246	sav16.0-pd7.0		Hu
RQSTDFGTT	Phage display	R Q > > D/E	savinase	savinase	R247 Q245 S240/S242	sav2.3	Ra	
FCTNNCELS	Phage display	N > > E L	savinase	savinase	T143 N173 N140 E136 L135	sav7.2	Ra	
FCTNNCELS	Phage display	N > > E L	savinase	savinase	N117 N116 E112 L111	sav7.1	Ra	
DFHVKYAAQ	Phage display		savinase	savinase		sav8.0	Ra	
VAQYKALPVVLENA	Protein fragments		Fel d I	savinase	L135 P168 V139 L111 E112 N116	sav12.0-pd8.0		Hu
AAYPDV	Protein fragments	A > > > > Y P >	a-amylase inhibitor	savinase	A215 Y214 P40 D41 V81	sav13.0-pd13.1-13.2		Hu
EQIFFTSGL	Phage display	D/E Q I F F T	savinase	savinase	E271 Q12 I8	sav5.3	Ra	
VDAAF	Protein fragments		Poa p IX	savinase	V203 D181 A179 A187 F189	sav15.0-pd12.0	Hu	
AVLRDC	Protein fragments		a-amylase inhibitor	savinase	A232 V234 L250 R247 D197	sav18.2-pd18.1-18.2		Hu
RAFRRNANW	Phage display	A R > A	savinase	savinase	A272/A273 R275 R19 N18 A15/A16	sav10.1	Ra	
CTARLRAGNACG	Phage display	A R > A	savinase	savinase	A15/A16 R19 L21 R275 A272 A273 N269	sav10.2	Ra	
TFHDAPALQ	Phage display		savinase	savinase	H39 D41 A74/A73 P86 A88 L90	sav4.0	Ra	
CTARVVALGVCG	Phage display	A R > A	savinase	savinase	R145 V147 V149 A151 L124/L126 G127	sav10.3	Ra	
GRFSNSKFK	Phage display	L > G R S	savinase	savinase	L148 G146 R145 S144/S141 N140	sav9.1-je4.0-lip5.1-5.2	Ra	
RRFANDHTR	Phage display	R/K R F > N	savinase	savinase	K27 R45 N43 D41 H39 T38/T213	sav6.2	Ra	
KRFANTEPA	Phage display	R/K R F > N	savinase	savinase	K251 R247 A174 N173	sav6.1	Ra-Mo	Mo
YKVSAL	Protein fragments		a-amylase inhibitor	savinase	Y91 K27 V26 S24 G23 L21	sav14.0-pd14.0		Hu
TGKYVS	Protein fragments		a-amylase inhibitor	savinase	S24 G25 K27 Y91 V93	sav17.0-pd17.1-17.2		Hu

Table 3: PD498 antibody binding peptide sequences, epitope patterns and epitope sequences.

Epitope pattern	Source					
	Fel d I	pd498	V198 A254 Q252 Y276 K239 A235 L233 P86	pd8.0		Hu
A>>>>YP>	a-amylase inhibitor	pd498	*3aA Y1/Y2 P-4/P-1 D-2 V81	pd13.2		Hu
	Poa p IX	pd498	S182 Y6 G7 P8 T13 P14 A15 A16	pd11.0	Hu	
>KL>>	Poa p IX	pd498	Y171 K136 L135 A108 Y113	pd4.4	Hu	
	a-amylase inhibitor	pd498	Y48/Y37 K46 *44aaV A43 L42	pd14.0		Hu
	Poa p IX	pd498	V196/V198 D197 A174/A176 A169 F163	pd12.0	Hu	
KQS	Poa p IX	pd498	A142 A147 V148 K120 Q27 S24/S25	pd2.3	Hu	
KQS	pd498	pd498	R44 K89 Q27 S236 K120 G146	pd2.2	Ra	
	Der p II	pd498	*28aV T88 *44a K R44 A43 L42	pd7.0		Hu
>KL>>	pd498	pd498	N56/N55 K46 L91 A29/A119 T28	pd4.3	Ra	
>KL>>	pd498	pd498	N240/N243 K239 L233/L234 A16 T21 R22	pd4.1	Ra	
>KL>>	Poa p IX	pd498	Y37 K46 L91 A114 Y113	pd4.5	Hu	
>KL>>	pd498	pd498	N240/N243 K239 L233/L234 A16 T21 R22	pd4.1	Ra	
YI>KL	pd498	pd498	Y113 I111 A108/A138 K136 L135	pd3.1	Ra	
KQS	pd498	pd498	A115 K145 N243 N240 K239 Q237 S236	pd2.1	Ra	
>RY>K/R	pd498	pd498	R94 R53 Y48 Q117 R112 S109/S137	pd1.5-lac2.0	Ra	
	Phl p V	pd498	A169 Q167 F163 T162 S160 G193	pd10.0		Hu
YI>KL	pd498	pd498	Y276 I246 K239 L234 S236	pd3.2	Ra	
>KL>>	pd498	pd498	N240/N243 K239 L233/L234 R22 P86	pd4.2	Ra	
A>>>>YP>	a-amylase inhibitor	pd498	*3aA Y2 P14 D18 V19	pd13.1		Hu
KQS	Poa p IX	pd498	A15 A16 V274 K239 Q237 S236	pd2.4	Hu	
	a-amylase inhibitor	pd498	G146 K145 Y141 V139 S137	pd17.2		Hu
	a-amylase inhibitor	pd498	A273 V274 L233 R22 D87	pd18.1		Hu
AR>A	Par j 1 + Par o 1	pd498	N10 S12 A15/A16 R275 A273/A249 R247 A174 D197 S170	pd9.0	Hu+Ra	Hu
	pd498	pd498	R22 G23 L233 S236	pd6.2	Ra	
>RY>K/R	pd498	pd498	R94 R53 Y48 P57 K46 L91	pd1.4-lac2.0	Ra	
>RY>K/R	pd498	pd498	R94 R53 Y48 P57 K46 L91	pd1.4-lac2.0	Ra	
	a-amylase inhibitor	pd498	L96 R94 S33 V35 Y37	pd15.0		Hu
>RY>K/R	pd498	pd498	S109/S137 R112 Y141 N144 K145	pd1.3-lac2.0	Ra	
>RY>K/R	pd498	pd498	T162 R161 Y192 N191 K186	pd1.2-lac2.0	Ra	
>RY>K/R	pd498	pd498	T133/T134 R112 Y141 N144 K145	pd1.1-lac2.0	Ra	
	a-amylase inhibitor	pd498	A92 *44aaV L42 R44 D75	pd18.2		Hu
	a-amylase inhibitor	pd498	S236 G238 K239 Y276 V274 S270	pd17.1		Hu
	a-amylase inhibitor	pd498	S12 P14 W17 S-5 W-6	pd16.0		Hu
>RSA	pd498	pd498	S137 R112 S109 A108	pd5.0-lac1.0-lip4.0-sav3.1-3.2	Ra	
	pd498	pd498	S215 M217 I205 M222 G219	pd6.1	Ra	

Table 4: Lipolase antibody binding peptide sequences, epitope patterns and epitope sequences.

QRPPRYELE	Phage display	R P P R	lipolase	lipolase		lip1.0	Ra	
ELEYRPPRQ	Phage display	> E Y	lipolase	lipolase	L124 E129 Y164	lip2.1	Ra	
HEYDMRVAW	Phage display	> E Y	lipolase	lipolase	H215 E219 Y220	lip2.2	Ra	
HEYPMDFHL	Phage display	> E Y	lipolase	lipolase	H215 E219 Y220	lip2.2	Ra	
SEYSMSITP	Phage display	> E Y	lipolase	lipolase	S217 E219 Y220	lip2.3	Ra	
CVWPAHAPLSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
CSWPSPAPLSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
CDFPLHAPLSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
CLFPSPAPRSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
CDGPAPAPWSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
CSFPLPAPRSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
CVYPSPAPWSCG	Phage display	> P > > P A P > S	lipolase	lipolase	P253 P250 A243 P208/P207 S214/S216/S217	lip3.0	Ra	
PEYTMNALS	Phage display	> E Y	lipolase	lipolase	P218 E219 Y220	lip2.4	Ra	
CSRSAKGARLCG	Phage display	> R S A	lipolase	lipolase	R209 S214 A182	lip4.0-lac1.0-pd5.0-sav3.1-3.2	Ra	
LEYPMSASQ	Phage display	> E Y	lipolase	lipolase	L124 E129 Y164	lip2.1	Ra	
RKLTLSGRS	Phage display	L > G R S	lipolase	lipolase	L67 G65 R81 S83/S85	lip5.1- je4.0-sav9.0	Ra	
RKLTLSGRS	Phage display	L > G R S	lipolase	lipolase	L96/L97 G212 R209/R179 S214	lip5.2- je4.0-sav9.0	Ra	
SYGAPATPAA	Protein fragments		Poa p IX	lipolase	S170 Y171 G172 A173 P174 A150 T153	lip6.0	Hu	
PAAGYTPAAP	Protein fragments		Poa p IX	lipolase	A18/A19/A20 G65 Y53 T123	lip7.0	Hu	Hu
YKLAY	Protein fragments		Poa p IX	lipolase	Y138 K74 L75 A68 Y16	lip8.1	Hu	
YKLAY	Protein fragments		Poa p IX	lipolase	Y53 K127 L67 A68 Y16	lip8.2	Hu	
KYDDYVATLS	Protein fragments		Poa p IX	lipolase	Y194 D167 D165 Y164 V132 A131 L52 S54	lip9.0	Hu	
EVKATPAGEL	Protein fragments		Poa p IX	lipolase	E43 V44 K46 A47 T72	lip10.0	Hu	
CGYSNAQGVDYWI	Protein fragments		Der p I	lipolase	Y53 S54 N25/N26 A18/A19/A20 Q15 V44	lip15.0	Hu	Hu
VPGIDPNACHYMKC	Protein fragments		Der p II	lipolase	P256 I255 D254 P253 N200 H198 Y261	lip16.0		Hu
SPVTKRASLKIDSKK	Protein fragments		Der p II	lipolase	R179 A182 S216/S217 I238 K237 I235 D234 S224 K223	lip17.0		Hu
IMSALAMVYLGAK	Protein fragments		Ovalbumin	lipolase	V140 Y138 L69 A49 A47 K46	lip18.0	Hu	
ELGVRE	Protein fragments		a-amylase inhibitor	lipolase	E99 L97 G109/G177 V176 R175 D242	lip11.0		Hu
GCRKEV	Protein fragments		a-amylase inhibitor	lipolase	G106 C107 R108 K98 E99	lip12.0		Hu
LRSVYQ	Protein fragments		a-amylase inhibitor	lipolase	L147 R81 S79 V77 Y16 Q15	lip13.0		Hu
SGPWSW	Protein fragments		a-amylase inhibitor	lipolase	S170 G172 P174 W89 S83	lip14.0		Hu

Table 5: Amylase (Natalase) antibody binding peptide sequences, epitope patterns and epitope sequences.

ARIDPRGPS	Phage display	A > I D P R/K	amylase	amylase	A380 K381 I382 D383 P384 R389	je1.1	Ra	
ARIDPRHGS	Phage display	A > I D P R/K	amylase	amylase	A380 K381 I382 D383 P384 R389	je1.1	Ra	
CSVAKIDPRTCG	Phage display	A > I D P R/K	amylase	amylase	A109 K138 D140 P142 R144	je1.2	Ra	
CSVAKIDPRTCG	Phage display	A > I D P R/K	amylase	amylase	A380 K381 I382 D383 P384 R389	je1.1	Ra	
AKIDPKPDT	Phage display	A > I D P R/K	amylase	amylase	A109 K138 D140 P142 R144	je1.2	Ra	
AKIDPKPDT	Phage display	A > I D P R/K	amylase	amylase	A380 K381 I382 D383 P384 R389	je1.1	Ra	
ARIDPRHGS	Phage display	A > I D P R/K	amylase	amylase	A109 K138 D140 P142 R144	je1.2	Ra	
QIYNDTGPT	Phage display	Q > Y > D >	amylase	amylase	Q390 L386 Y368/Y367 D366	je2.4	Ra	
QIYNDTGPT	Phage display	Q > Y > D >	amylase	amylase	Q170 I173 Y196 D195	je2.3	Ra	
QIYNDTGPT	Phage display	Q > Y > D >	amylase	amylase	Q357 I352 Y349 D366	je2.2	Ra	
QIYNDTGPT	Phage display	Q > Y > D >	amylase	amylase	Q331 I370 Y368/Y367 D366	je2.1	Ra	
CGSATIDPRQCG	Phage display	A > I D P R/K	amylase	amylase	A109 K138 D140 P142 R144	je1.2	Ra	
CNADNQMYPQCG	Phage display	A > > > > Y P >	amylase	amylase	N29 A27 D26/D25 Y8 P41/P42	je3.1	Ra	
ARIDPRGPS	Phage display	A > I D P R/K	amylase	amylase	A109 K138 D140 P142 R144	je1.2	Ra	
CGSATIDPRQCG	Phage display	A > I D P R/K	amylase	amylase	A380 K381 I382 D383 P384 R389	je1.1	Ra	
CDADSSGYPLCG	Phage display	A > > > > Y P >	amylase	amylase	A107/A109 D108 Y57 P41/42	je3.3	Ra	
QLYGDEQLP	Phage display	Q > Y > D >	amylase	amylase	Q331 I370 Y368/Y367 D366	je2.1	Ra	
QLYGDEQLP	Phage display	Q > Y > D >	amylase	amylase	Q357 I352 Y349 D366	je2.2	Ra	
QLYGDEQLP	Phage display	Q > Y > D >	amylase	amylase	Q170 I173 Y196 D195	je2.3	Ra	
QLYGDEQLP	Phage display	Q > Y > D >	amylase	amylase	Q390 L386 Y368/Y367 D366	je2.4	Ra	
RYAQIDPRW	Phage display	A > I D P R/K	amylase	amylase	A380 K381 I382 D383 P384 R389	je1.1	Ra	
RYAQIDPRW	Phage display	A > I D P R/K	amylase	amylase	A109 K138 D140 P142 R144	je1.2	Ra	
GEFNLGRSS	Phage display	L > G R S	amylase	amylase	L88 G92 R31 S28	je4.1-sav9.0-lip5.1-5.2	Ra	
CNADSWGYPRCG	Phage display	A > > > > Y P >	amylase	amylase	N29 A27 D26/D25 Y8 P41/P42	je3.1	Ra	
CNADNQMYPQCG	Phage display	A > > > > Y P >	amylase	amylase	N102 A233 D232 Y54 P41/P42	je3.2	Ra	
CNADSWGYPRCG	Phage display	A > > > > Y P >	amylase	amylase	N102 A233 D232 Y54 P41/P42	je3.2	Ra	
GEFNLGRSS	Phage display	L > G R S	amylase	amylase	L62 G63/G76 R78 S79	je4.2-sav9.0-lip5.1-5.2	Ra	

Table 6: Cellulase (Carezyme; Cel45 from Humicola insolens) antibody binding peptide sequences, epitope patterns and epitope sequences.

CVHAGPRAGTCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
CVHAGPRAGTCG	Phage display	V H > G >	carezyme	carezyme		car2.0	Ra	
CLSGPLAGRVCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
CRISPWYSVPCG	Phage display		carezyme	carezyme		car3.0	Ra	
CLSGPAAGQSCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
	Phage display	A > I D P R/K	je-1	carezyme	R146 I131 D133 P137	car11.2	Ra	
CITRGTRAGWCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
	Phage display	A R > A	savinase	carezyme	A191 R200 R201 A83 N81	car6.2	Ra	
CLSGPAAGQSCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
	Phage display	A > I D P R/K	je-1	carezyme	A195 R37 I38 D40 L44	car11.1	Ra	
	Phage display	Q > Y > D >	savinase, je-1	carezyme	Q59 Y54 G134 D133 T136	car10.0	Ra	
	Phage display	> P > > A P > S	lipoprime	carezyme	W62/W169 P61 P165 A162 P160	car9.0	Ra	
CITRGTRAGWCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
	Phage display	R/K R F > N	savinase	carezyme	R7 R170 F174 A177	car7.0	Ra	
CLSGPLAGRVCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
	Phage display	A R > A	savinase	carezyme	A1 R4 R7 A177 N176	car6.1	Ra	
	Phage display	> P > R D T G	laccase	carezyme	D178 P180 R4 D2 S183	car5.0	Ra	
	Phage display	> R Y > K/R	pd498	carezyme	R170 R153 Y168 P165 K164 L163	car4.0	Ra	
	Phage display	D/E Q I F F T	savinase	carezyme	Q36 I38 F41 F29 T197	car8.0	Ra	
CLTAGPSAGYCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
CYTTGRLAGLCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
CYTTGRLAGLCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
CVHSGPRAGYCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
CVHSGPRAGYCG	Phage display	V H > G >	carezyme	carezyme		car2.0	Ra	
CVHAGPRAGTCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
CVHSGPRAGYCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
CVHSGLSRRLLR	Phage display	V H > G >	carezyme	carezyme		car2.0	Ra	
CVTRGPNAGSCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
CLTAGPSAGYCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
CVTRGPNAGSCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	
CITSGPRAGNCG	Phage display	> G > > A G	carezyme	carezyme	T95/S96 G27 P98 A100 G101	car1.2	Ra	
CITSGPRAGNCG	Phage display	> G > > A G	carezyme	carezyme	P23 R201 A83 G84	car1.1	Ra	

Table 7: Laccase (*Myceliophtora thermopila* laccase) antibody binding peptide sequences, epitope patterns and epitope sequences.

PQSDPGESQ	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
WPKSDAGDS	Phage display	P > > D A G	laccase	laccase	P241 R409 S410/S416 D434 A389 G390	lac4.1	Ra	
PQSDAGVVM	Phage display	P > > D A G	laccase	laccase	P241 R409 S410/S416 D434 A389 G390	lac4.1	Ra	
DPVRDTGAG	Phage display	> P > R D T G	laccase	laccase	P241 R409 D434 T432 G430/G390	lac5.1	Ra	
GPSRDAGLL	Phage display	P > > D A G	laccase	laccase	P241 R409 S410/S416 D434 A389 G390	lac4.1	Ra	
PASDAGRGP	Phage display	P > > D A G	laccase	laccase	P241 R409 S410/S416 D434 A389 G390	lac4.1	Ra	
PRDSTGLAL	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
PQSDPGESQ	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
RYPFLRATN	Phage display	> R Y > K/R	laccase	laccase		lac2.0-pd1.1-1.4	Ra	
GAARDARSA	Phage display	> R S A	laccase	laccase		lac1.0-lip4.0-pd5.0-sav3.1-3.2	Ra	
PRSDTGFGS	Phage display	> P > R D T G	laccase	laccase	P241 R409 D434 T432 G430/G390	lac5.1	Ra	
LPRSDPGGR	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
DPARDTGDV	Phage display	> P > R D T G	laccase	laccase	P241 R409 D434 T432 G430/G390	lac5.1	Ra	
APKSDNGIT	Phage display	P > > D A G	laccase	laccase	P241 R409 S410/S416 D434 A389 G390	lac4.1	Ra	
PKSDPGTNW	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
PRTDPGWLA	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
LPRSDPGGR	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
PSSDPGARS	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
HVFDKNVTR	Phage display		laccase	laccase		lac6.0		
PRSDPGTPT	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
PRSDPGTPT	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
PRDSTGLAL	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
PRTDPGWLA	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
PSSDPGARS	Phage display	P > S/T D P G	laccase	laccase	P378 R379 T442 D443 P445 G446	lac3.1	Ra	
PKSDPGTNW	Phage display	P > S/T D P G	laccase	laccase	P180 R175 T168 D166 P165 G265	lac3.2	Ra	
WPKSDAGDS	Phage display	P > > D A G	laccase	laccase	P350 S349 D80 A79 G78	lac4.2	Ra	
PQSDAGVVM	Phage display	P > > D A G	laccase	laccase	P350 S349 D80 A79 G78	lac4.2	Ra	

GPSRDAGLL	Phage display	P > > D A G	laccase	laccase	P350 S349 D80 A79 G78	lac4.2	Ra	
PASDAGRGP	Phage display	P > > D A G	laccase	laccase	P350 S349 D80 A79 G78	lac4.2	Ra	
APKSDNGIT	Phage display	P > > D A G	laccase	laccase	P350 S349 D80 A79 G78	lac4.2	Ra	
WPKSDAGDS	Phage display	P > > D A G	laccase	laccase	P300 R234 S211 D213 A296	lac4.3	Ra	
PQSDAGVVM	Phage display	P > > D A G	laccase	laccase	P300 R234 S211 D213 A296	lac4.3	Ra	
GPSRDAGLL	Phage display	P > > D A G	laccase	laccase	P300 R234 S211 D213 A296	lac4.3	Ra	
PASDAGRGP	Phage display	P > > D A G	laccase	laccase	P300 R234 S211 D213 A296	lac4.3	Ra	
APKSDNGIT	Phage display	P > > D A G	laccase	laccase	P300 R234 S211 D213 A296	lac4.3	Ra	
DPVRDTGAG	Phage display	> P > R D T G	laccase	laccase	P378 R379 D469 T473 G446	lac5.2	Ra	
PRSDTGFGS	Phage display	> P > R D T G	laccase	laccase	P378 R379 D469 T473 G446	lac5.2	Ra	
DPARDTGDV	Phage display	> P > R D T G	laccase	laccase	P378 R379 D469 T473 G446	lac5.2	Ra	
DPVRDTGAG	Phage display	> P > R D T G	laccase	laccase	P60 R59 D51/D53 T10/T12 G30	lac5.3	Ra	
PRSDTGFGS	Phage display	> P > R D T G	laccase	laccase	P60 R59 D51/D53 T10/T12 G30	lac5.3	Ra	
DPARDTGDV	Phage display	> P > R D T G	laccase	laccase	P60 R59 D51/D53 T10/T12 G30	lac5.3	Ra	
DPVRDTGAG	Phage display	> P > R D T G	laccase	laccase	P157/P155 R23 D118 T114 G113	lac5.4	Ra	
PRSDTGFGS	Phage display	> P > R D T G	laccase	laccase	P157/P155 R23 D118 T114 G113	lac5.4	Ra	
DPARDTGDV	Phage display	> P > R D T G	laccase	laccase	P157/P155 R23 D118 T114 G113	lac5.4	Ra	

Birch allergen:

Bet v1 (WO99/47680) was used as the parent protein for identification of epitope sequences that may cross react with enzyme epitopes. The structural coordinates from 1BV1.pdb (Gajhede et al., NAT.STRUCT.BIOL., Vol. 3, p. 1040, 1996) were used as well the corresponding sequence (Swissprot accession number P15494). The epitope pattern P>PAP>S (which had been identified from antibody binding peptides specific for anti-Lipolase antibodies) was found to match three (overlapping) epitope sequences on the surface of Bet v1:

Bet v1 1.1: P31 A34 P35 A37 P59 S39/S40;

Bet v1 1.2: P63 L62 P59 A37 P35 S39/S40; and

Bet v1 1.3: P59 S39/S40 P31 A34 P35 S39/S40.

Example 2

Localisation of epitope sequences and epitope areas on the 3D-structure of acceptor proteins.

Epitope sequences were assessed on the 3D-structure of the protein of interest, using apropriate software (e.g. SwissProt Pdb Viewer, WebLite Viewer).

In a first step, the identified epitope patterns were fitted with the 3D-structure of the enzymes. A sequence of at least 3 amino acids, defining a specific epitope pattern, was localised on the 3D-structure of the acceptor protein. Conservative mutations (e.g. aspartate for glutamate, lysine for arginine, serine for threonine) were considered as one for those patterns for which phage display had evidenced such exchanges to occur. Among the possible sequences provided by the protein

structure, only those were retained where the sequence matched a primary sequence, or where it matched a structural sequence of amino acids, where each amino acid was situated within a distance of 5Å from the next one. Occasionally, the mobility of the amino acid side chains, as provided by the software programme, had to be taken in to consideration for this criterium to be fulfilled.

Secondly, the remaining anchor amino acids as well as the variable amino acids, i.e. amino acids that were not defining a pattern but were present in the individual sequences identified by phage library screening, were assessed in the area around the various amino acid sequences localised in step 1. Only amino acids situated within a distance of 5Å from the next one were included.

Finally, an accessibility criterium was introduced. The criterium was that at least half of the anchor amino acids had a surface that was >30% accessible. Typically, 0-2 epitopes were retained for each epitope pattern. In some cases, two different amino acids could with equal probability be part of the epitope (e.g. two leucines located close to each other in the protein 3D-structure). For example, in Savinase two epitopes actually fit to the antibody binding peptide LDQIFFTRW: L75 D41 Q2 I79 and L42 D41 Q2 I79. A shorthand notation for such a situation is: L42/L75 D41 Q2 I79.

Thus, a number of epitope sequences were identified and localised on the surface of various proteins. As suggested by sequence alignment of the antibody binding peptides, structural analysis confirmed most of the epitopes to be enzyme specific, with only few exceptions. Overall, most of the identified epi-

topes were at least partially structural. However, some proteins (e.g. amylase) expressed predominantly primary sequence epitopes. Typically, the epitopes were localised in very discrete areas of the enzymes, and different epitope sequences often shared some amino acids (hot-spots).

The identified epitope sequences are shown in Tables 2-7.

Example 3

Epitope areas

It is common knowledge that amino acids that surround binding sequences can affect binding of a ligand without participating actively in the binding process. Based on this knowledge, areas covered by amino acids with potential steric effects on the epitope-antibody interaction, were defined around the identified epitopes. Practically, all amino acids situated within 5Å from the amino acids defining the epitope were included. The accessibility criterium was not included for defining epitope areas, as hidden amino acids can have an effect on the surrounding structures.

For Savinase, the following amino acid residues belong to the epitope area that correspond to each epitope sequence indicated in Table 2:

sav1.1	A1	Q2	S3	P5	H39	P40	D41	L42	N43
	G63	T66	H67	A69	G70	T71	A73	A74	L75	N77
	S78	I79	G80	V81	L82	G83	N204	V205	Q206	

S207 T208 Y209 P210 S212 T213 Y214 A215 S216
L217

sav1.2 S153 G154 N155 S156 G157 A158 G160 S161
I162 S163 A169 R170 A174 M175 A176 V177 G178
R186 F189 S190 Q191 Y192 G193 A194 G195 L196
D197 I198 V199 T220 R247 K251 A254 T255 S256
T260 N261 L262 Y263 G264 S265 G266 L267

sav2.1 W6 G7 I8 R10 V11 Q12 A13 P14 A15
A16 R19 L21 V84 T180 D181 Q182 N183 N184
I198 V199 A200 P201 H226 V227 A230 L233 V234
K237 N238 H249 L250 T253 A254 T255 S256 L257
S265 G266 L267 V268 N269 A270 E271 A272 A273
T274 R275

sav2.2 S153 G154 N155 S156 G157 A158 S161 I162
S163 G178 A179 T180 D181 N184 N185 R186 A187
S188 F189 S190 Q191 Y192 G193 L196 T220 L262
Y263

sav2.3 A142 T143 G146 V147 L148 Y171 A172 N173
A174 M175 D197 A231 V234 K235 N238 P239 S240
W241 S242 N243 V244 Q245 I246 R247 N248 H249
L250 K251

sav3.1 S153 G154 N155 S156 G157 A158 V177 G178
A179 T180 D181 N184 N185 R186 A187 S188 F189
S190 Q191 Y192 V199 A200 P201 G202 V203 N218
G219 T220 A223 L262 Y263

sav3.2	L111	E112	G115	N116	M119	A138	V139	N140
S141	A142	S144	R145	G146	V147	V149	N173	N243

sav4.0	Q2	H17	T22	G23	S24	G25	V26	K27	V28
V30	I35	S37	T38	H39	P40	D41	L42	N43	I44
R45	G46	T66	A69	G70	T71	I72	A73	A74	L75
N76	N77	I79	G80	V81	L82	G83	V84	A85	P86
S87	A88	E89	L90	Y91	A92	T208	Y209	P210	
S212	T213	Y214							

sav5.1	A1	Q2	S3	V4	I35	S37	H39	P40	D41
L42	N43	I44	T66	A69	G70	A73	A74	L75	N76
N77	S78	I79	G80	V81	L82	G83	P86	L90	
T208	Y214								

sav5.2	V30	T33	G34	I35	S37	T38	L42	N43	I44
R45	G46	E54	S57	T58	Q59	D60	G61	N62	G63
H64	G65	T66	H67	A69	L90	Y91	A92	K94	
P210									

sav5.3	V4	P5	W6	G7	I8	S9	R10	V11	Q12
A13	P14	A15	A16	R19	N269	A270	E271	A272	
A273	T274	R275							

sav5.4	A1	Q2	P40	D41	F50	L75	N77	S78	I79
G80	V81	V104	S105	S106	I107	A108	Q109	G110	
L111	E112	W113	A114	G115	N116	Q137	A138	S141	
A142	Y214								

sav6.1	V139	N140	T143	L148	V149	A151	P168	A169
Y171	A172	N173	A174	M175	A176	D197	I198	N243
V244	Q245	I246	R247	N248	H249	L250	K251	N252

T253 A254 S265

sav6.2 Q2 G25 V26 K27 V28 A29 I35 S37 T38
H39 P40 D41 L42 N43 I44 R45 G46 G47 Q59
T66 A69 G70 A73 A74 L75 N77 I79 G80 V81
L82 A88 E89 L90 Y91 N117 G118 M119 H120
V121 S207 T208 Y209 P210 G211 S212 T213 Y214
A215

sav7.1 K27 L31 I107 A108 Q109 G110 L111 E112
W113 A114 G115 N116 N117 G118 M119 A122 L124
L135 Q137 A138 V139 S141 A142 R145 V149

sav7.2 V104 I107 A108 L111 S132 A133 T134 L135
E136 Q137 A138 V139 N140 S141 A142 T143 S144
R145 G146 V147 V149 Y167 P168 Y171 A172 N173
A174 M175 N243 R247

sav9.1 L111 E112 A114 G115 N116 M119 H120 V121
A122 E136 Q137 A138 V139 N140 S141 A142 T143
S144 R145 G146 V147 L148 V149 V150 N173 M175
N243 I246 R247 L250

sav9.2 L126 G127 S128 P129 A152 S153 G154 S161
I162 S163 Y167 P168 A169 R170 Y171 A172 A176
V177 G178 Q191 Y192 G193 A194 G195 L196 D197
I198 V199 T260 N261 L262 Y263 G264

sav10.1 Q12 A13 P14 A15 A16 H17 N18 R19 G20
L21 T22 N76 L82 G83 V84 A85 P86 L233
V234 K237 N238 H249 L250 T253 N269 A270 E271

A272 A273 T274 R275

sav10.2 V11 Q12 A13 P14 A15 A16 H17 N18 R19
G20 L21 T22 G23 L233 V234 Q236 K237 N238
H249 L250 T253 A254 T255 L267 V268 N269 A270
E271 A272 A273 T274 R275

sav10.3 L31 D32 H64 V68 V95 L96 I107 L111
A114 G115 N116 M119 V121 A122 N123 L124 S125
L126 G127 S128 P129 V139 S141 A142 T143 S144
R145 G146 V147 L148 V149 V150 A151 A152 S153
S163 Y167 P168 A169 N173 A174 M175 A176 V177
T220 S221 M222 T224 P225 V227 A228 A231 N243
I246 R247 L250

sav10.4 P131 S132 A133 L135 E136 V139 A151 A152
S153 G160 S161 I162 S163 Y167 P168 A169 R170
Y171 A172 N173 A174 A176 Q191 Y192 G193 A194
G195 L196 R247 S259 T260 N261 L262 Y263 G264

sav11.0 W6 G154 N155 S156 G157 A179 T180 D181
Q182 N183 N184 N185 R186 A187 S188 F189 S190
Q191 Y192 P201 G202 V203 N204 V205 L217 N218
G219 T220 L262 Y263

sav12.0 L31 I107 A108 Q109 G110 L111 E112 W113
A114 G115 N116 N117 G118 A122 L124 S132 A133
T134 L135 Q137 A138 V139 N140 S141 T143 R145
V149 A151 S163 Y167 P168 A169 R170 Y171 N173
A174

sav13.0 Q2 S3 P5 T38 H39 P40 D41 L42 N43
H67 G70 A73 A74 L75 N77 I79 G80 V81 L82

	G83	V205	Q206	S207	T208	Y209	S212	T213	Y214
	A215	S216	L217						
sav14.0	A16	H17	R19	G20	L21	T22	G23	S24	G25
	V26	K27	V28	A29	V30	I35	I44	R45	G46	G47
	V84	A85	P86	S87	A88	E89	L90	Y91	A92	V93
	W113	N117	G118	M119	H120	V121	A232	L233	K235	
	Q236	K237	T274						
sav15.0	W6	R10	G154	N155	S156	G157	V177	G178	
	A179	T180	D181	Q182	N183	N184	N185	R186	A187	
	S188	F189	S190	Q191	V199	A200	P201	G202	V203	
	N218	G219	T220	A223	L257	Y263	L267			
sav16.0	A13	A16	H17	G20	L21	T22	G23	S24	G25
	V26	V28	I72	A73	V84	A85	P86	S87	A88	E89
	L90	H120	G229	A230	A231	A232	L233	V234	K235	
	Q236	K237	N238	P239	S240	W241	I246	H249	L250	
	A270	A273	T274						
sav17.0	T22	G23	S24	G25	V26	K27	V28	A29	V30
	L31	D32	I35	I44	R45	G46	G47	A48	F50	S87
	A88	E89	L90	Y91	A92	V93	K94	V95	G110	
	W113	N117	G118	M119	H120	V121	A232	K235	Q236	
sav18.1	W6	G7	I8	S9	R10	V11	Q12	A179		
	T180	D181	Q182	N183	N184	N185	R186	A187	I198	
	V199	A200	P201	V203	H226	V227	A230	H249	L250	
	K251	N252	T253	A254	T255	S256	L257	S265	G266	
	L267	V268	N269	A270						

sav18.2	A13	A16	H17	L21	T22	G23	V26	V28	V84
	A85	A88	V121	L148	Y171	A172	N173	V174	M175
	A176	G195	L196	D197	I198	V199	V227	A228	G229
	A230	A231	A232	L233	V234	K235	Q236	K237	N238
	W241	N243	V244	Q245	I246	R247	N248	H249	L250
	K251	N252	T253	A254	Y263	G264	S265	G266	V268
	A270	A273	T274						

sav19.1	A16	H17	R19	G20	L21	T22	G23	S24	G25
	V26	K27	V28	S87	A88	E89	H120	V121	A232
	L233	V234	K235	Q236	K237	N238	P239	T274	

sav19.2	A1	Q2	S3	V4	P5	D41	H64	H67	G70
	T71	A74	L75	N77	S78	I79	G80	V81	L82	G83
	G202	V203	N204	V205	Q206	S207	T208	Y209	Y214
	A215	S216	L217	N218	G219	M222			

For PD498, the following amino acid residues belong to the epitope area that correspond to each epitope sequence indicated in Table 3:

pd1.1 D105	A108	S109	G110	I111	R112	Y113	A114	A115
D116	Q117	N131	S132	T133	T134	L135	K136	S137
A138	V139	D140	Y141	A142	W143	N144	K145	G146
A147								

pd1.2 C128	E129	A153	G154	N155	D156	N157	V158	S160
R161	T162	F163	Q167	S170	G178	A179	I180	D181
D184	R185	K186	A187	S188	F189	S190	N191	Y192

G193 T194 W195 V196 T220 T262 N263

pd1.3 F50 L104 D105 S106 I107 A108 S109 G110 I111
R112 Y113 A114 A115 D116 Q117 T133 T134 L135
K136 S137 A138 V139 D140 Y141 A142 W143 N144
K145 G146 A147

pd1.4 T28 *28aV A29 V30 D32 S33 G34 V35 Y37
*44aaV I45 K46 G47 Y48 D49 F50 I51 R53
D54 N55 N56 P57 M58 D60 L61 K89 I90 L91
A92 V93 R94 V95 L96 D97 A98 Y113 A114
Q117 A119

pd1.5 D32 S33 G34 K46 G47 Y48 D49 F50 I51 D52
R53 D54 N55 P57 M58 L61 L91 A92 V93 R94
V95 L96 D97 A98 L104 D105 S106 I107 A108
S109 G110 I111 R112 Y113 A114 A115 D116 Q117
G118 A119 T133 T134 L135 K136 S137 A138 V139
D140 Y141 A142

pd2.1 V19 T21 I111 R112 Y113 A114 A115 D116 Q117
G118 A119 L122 D140 Y141 A142 W143 N144 K145
G146 A147 V148 L233 L234 A235 S236 Q237 G238
K239 N240 N243 V244 Q245 I246 R247 Q248 A249
A273 V274 R275 Y276

pd2.2 S24 S25 T26 Q27 T28 *28aV L42 A43 R44
*44aK *44aaV I45 D75 N77 D87 T88 K89 I90

L91 G118 A119 K120 V121 L122 G146 A147 V148
A232 A235 S236

pd2.3 R22 G23 S24 S25 T26 Q27 T28 *28aV D87 T88
K89 I111 A115 G118 A119 K120 V121 L122 S137
A138 V139 D140 Y141 A142 W143 N144 K145 G146
A147 V148 V149 V150 I175 A231 A232 A235 S236
N243 I246 R247

pd2.4 W-6 S12 T13 P14 A15 A16 V19 T21 R22 G23
S24 Q27 L230 A231 L233 L234 A235 S236 Q237
G238 K239 N240 N243 Q245 I246 S270 N271 K272
A273 V274 R275 Y276

pd3.1 L31 K46 G47 Y48 F50 L91 V93 S103 L104
D105 S106 I107 A108 S109 G110 I111 R112 Y113
A114 A115 D116 Q117 G118 L122 L124 C130 S132
T133 T134 L135 K136 S137 A138 V139 D140 Y141
A142 Q167 P168 Y171 P172

pd3.2 V19 T21 R22 G23 S24 Q27 K120 V121 V148
L230 A231 A232 L233 L234 A235 S236 Q237 G238
K239 N240 N243 Q245 I246 R247 Q248 A249 I250
Q252 T253 K272 A273 V274 R275 Y276

pd4.1 W-6 S12 T13 P14 A15 A16 W17 D18 V19 T21
R22 G23 S24 M84 A85 P86 D87 T88 A142

W143 G146 A147 V148 G229 L230 A231 A232 L233
L234 A235 S236 Q237 G238 K239 N240 N243 V244
Q245 I246 R247 Q248 A249 I250 S270 N271 A273
V274 R275 Y276

pd4.2 W-6 T13 A16 W17 V19 T21 R22 G23 S24
*44aK A73 A74 *75aT G83 M84 A85 P86 D87 T88
A142 G146 G146 A147 V148 G229 L230 A231 A232
L233 L234 A235 S236 Q237 G232 K239 N240 N243
V244 Q245 I246 R247 Q248 A249 I250 S270 A273
V274 R275 Y276

pd4.3 T26 Q27 T28 *28aV A29 V30 L31 Y37 *44aaV
I45 K46 G47 Y48 D49 D52 R53 D54 N55 N56
P57 M58 V72 T88 K89 I90 L91 A92 V93
Y113 A114 A115 Q117 G118 A119 K120 V121 L122
N123 A147 A228 A232

pd4.4 K46 G47 F50 L91 V93 S103 L104 D105 S106
I107 A108 S109 G110 I111 R112 Y113 A114 A115
D116 Q117 G118 C130 S132 T133 T134 L135 K136
S137 A138 V139 D140 Y141 Q167 P168 A169 S170
Y171 P172 N173 A174

pd4.5 T28 *28aV A29 V30 L31 V35 D36 Y37 N38 H39
L42 A43 *44aaV I45 K46 G47 Y48 F50 N55
N56 P57 M58 K89 I90 L91 A92 V93 A108
S109 G110 I111 R112 Y113 A114 A115 D116 Q117
G118 A119 L122

pd5.0	F50	S103	L104	D105	S106	I107	A108	S109	G110
	I111	R112	Y113	A114	A115	D116	Q117	T133	T134
	L135	K136	S137	A138	V139	D140	Y141	A142	

pd6.1	Y4	Y6	G7	G63	H64	H67	V68	T71	N155
	A179	F189	P201	G202	V203	N204	I205	A206	S207
	V209	G213	Y214	S215	Y216	M217	S218	G219	T220
	S221	M222	A223	S224	P225	H226			

pd6.2	W-6	T13	A16	W17	V19	T21	R22	G23	S24	S25
	Q27	M84	A85	P86	D87	T88	G229	L230	A231	
	A232	L233	L234	A235	S236	Q237	G238	S270	V274	

pd7.0	R22	G23	S24	S25	Q27	T28	*28aV	A29	V30	V35
	D36	Y37	N38	H39	P40	D41	L42	A43	R44	
	*44aK	*44aaV		T66	A69	G70	V72	A73	A74	D75
	N77	A85	P86	D87	T88	K89	I90	L91	A119	
	V121	L122	N123	T208	A228	A231				

pd8.0	W-6	T13	A16	W17	T21	R22	G23	Q27	*44aK	A73
	A74	*75aT	G83	M84	A85	P86	D87	T88	K120	
	V121	I175	A176	V177	G178	V196	D197	V198	T199	
	A200	V227	G229	L230	A231	A232	L233	L234	A235	
	S236	Q237	G238	K239	N240	N243	Q245	I246	Q248	
	A249	I250	Q252	T253	A254	F264	Y265	G266	I268	

pd9.0	W-6	Y6	G7	P8	Q9	N10	T11	S12	T13	P14
	A15	A16	W17	D18	V19	T21	M84	V139	W143	
	V148	V149	A151	P168	A169	Y171	P172	N173	A174	
	I175	A176	D181	S182	N183	D184	D197	P201	L230	
	L233	L234	K239	N240	N243	V244	Q245	I246	R247	
	Q248	A249	I250	E251	Q252	T253	A254	K267	I268	
	N269	S270	N271	K272	A273	V274	R275	Y276		

pd10.0	L124	L126	G127	C128	E129	C130	N131	L135	
V139	A151	A152	A153	G154	N155	D156	N157	V158	
S160	R161	T162	F163	Q167	P168	A169	S170	Y171	
A174	I175	A176	N191	Y192	G193	T194	W195	V196	
T262	N263	F264	*264aK						

pd11.0	W-6	S-5	Y2	Y4	Q5	Y6	G7	P8	Q9
N10	T11	S12	T13	P14	W17	D18	V19	T21	A82
M84	I180	D181	S182	N183	D184	P201	G202	V203	
N204	I205	H226	L233	S270	N271	V274	R275		

pd12.0	G127	C128	E129	V139	V148	V149	V150	A151
A152	A153	G154	N155	D156	V158	R161	T162	F163
Q167	P168	A169	S170	Y171	P172	N173	A174	I175
A176	V177	G178	N191	Y192	G193	T194	W195	V196
D197	V198	T199	A200	V227	R247	I250	E251	A254
N263	F264	*264aK		Y265	G266	I268		

pd13.1	W-6	S-5	P-4	D-2	P-1	Y1	Y2	S3	
*3aA	Y4	Q5	P8	Q9	S12	T13	P14	A15	A16
W17	D18	V19	T21	R22	G80	V81	A82	N271	

V274 R275

pd13.2	W-6	S-5	P-4	N-3	D-2	P-1	Y1	Y2	S3
*3aA	Y4	Q5	P8	Q9	P14	W17	D41	G70	A74
D75	*75aT	N76	N77	G78	I79	G80	V81	A82	G83
A206	S207	T208	Y214						

pd14.0	T28	V35	D36	Y37	N38	H39	P40	D41	L42
A43	R44	*44aK	*44aaV		I45	K46	G47	Y48	D49
F50	R53	D54	N55	N56	P57	M58	T66	A69	G70
A73	A74	D75	K89	I90	L91	A92	V93	R94	
Y113	T208								

pd15.0	V30	L31	D32	S33	G34	V35	D36	Y37	N38
H39	L42	A43	*44aaV		K46	Y48	D49	F50	I51
N56	P57	M58	D60	L61	N62	G63	H64	G65	T66
A69	I90	A92	V93	R94	V95	L96	D97	A98	
G100	S101	G102	S103	S106	I107	G110	S125	L126	
V209	P210	N211	N212						

pd16.0	W-6	S-5	P-4	N-3	Y2	G7	P8	Q9	N10
T11	S12	T13	P14	A15	A16	W17	D18	V19	T21
R22	*75aT	N76	A82	G83	M84	A85	P86	L233	
N269	S270	N271							

pd17.1	T11	S12	A15	A16	D18	V19	T21	R22	G23
S24	Q27	L230	A232	L233	L234	A235	S236	Q237	
G238	K239	N240	N243	Q245	I246	Q248	A249	Q252	

T253 N269 S270 N271 K272 A273 V274 R275 Y276

pd17.2 A108 I111 R112 A115 D116 K120 L124 T133
T134 L135 K136 S137 A138 V139 D140 Y141 A142
W143 N144 K145 G146 A147 V148 V149 P168 Y171
N173 A174 N243

pd18.1 W-6 T13 A16 W17 V19 T21 R22 G23 S24
S25 *44aK M84 A85 P86 D87 T88 K89 G229
L230 A231 A232 L233 L234 A235 S236 Q237 K239
A249 I250 T253 N269 S270 N271 K272 A273 V274
R275 Y276

pd18.2 D-2 V30 V35 D36 Y37 N38 H39 P40 D41
L42 A43 R44 *44aK *44aaV I45 K46 G47 Y48
P57 T66 A69 G70 A73 A74 D75 *75aT N76 N77
I79 V81 A82 A85 P86 D87 T88 K89 I90 L91
A92 V93 R94 T208

For Lipolase, the following amino acid residues belong to the epitope area that correspond to each epitope sequence indicated in Table 4:

lip2.1 Y53 F55 V63 L78 F80 W117 V120 A121
D122 T123 L124 R125 Q126 K127 V128 E129 D130
A131 V132 R133 V140 L159 R160 G161 N162 G163
Y164 D165 I166 G190

lip2.2 V2 L6 F10 A173 P174 R175 A182 L193
Y194 R195 I196 T197 P204 R205 Y213 S214 H215
S216 S217 P218 E219 Y220 W221 I222 I235 V236
K237 I238 E239 I241 D242 A243 G246 N247 N248

lip2.3 V2 L6 F10 A182 L185 T186 L193 Y194
R195 I196 T197 H215 S216 S217 P218 E219 Y220
W221 I222 I235 V236 K237 I238 E239 G240 I241
A243 G246 N247 N248

lip2.4 V2 L6 F10 L193 Y194 R195 I196 T197
S216 S217 P218 E219 Y220 W221 I222 I235 V236
K237 I238 E239 G240 A243 G246 N247 N248

lip3.0 L93 K94 F95 H110 A173 P174 R175 V176
G177 N178 R179 A182 L185 T186 L193 R195 N200
P201 I202 P204 R205 L206 P207 P208 R209 E210
F211 G212 Y213 S214 H215 S216 S217 P218 E219
I238 E239 G240 I241 D242 A243 T244 G245 N248
R259 P250 N251 I252 P253 D254 I255

lip4.0 R175 V176 G177 N178 R179 A180 F181 A182
E183 F184 L185 T186 R205 P207 P208 R209 E210
F211 G212 Y213 S214 H215 S216 S217 I241 D242
N248

lip5.1 A20 Y21 N25 N26 T50 F51 L52 Y53 S54
F55 E56 V63 T64 G65 F66 L67 A68 L69 I76
V77 L78 S79 F80 R81 G82 S83 R84 S85 I86
E87 N88 W89 K127 V128 A131 H145 S146 L147
G148 L151 G266

lip5.2 K94 F95 L96 L97 K98 E99 R108 G109
H110 D111 G112 R175 V176 G177 N178 R179 A180
F181 A182 E183 F184 R205 P207 P208 R209 E210
F211 G212 Y213 S214 H215 S216 I241 D242 N248

lip6.0 Q9 F10 N11 F13 A14 S17 V63 F80 R81
W89 L93 F113 S116 W117 F142 T143 G144 H145
S146 L147 G148 G149 A150 L151 A152 T153 V154
A155 G156 A157 V168 F169 S170 Y171 G172 A173
P174 R175 V176 F181 L185 L193 Y194 R195 I196
T197 D201 V203 P204 L206 P207 H215 H258 Y261
F262 I265

lip7.0 F13 A14 Q15 Y16 S17 A180 A19 A20 Y21
C22 G23 N25 N26 I34 C36 A40 C41 F51 L52
Y53 S54 F55 E56 V63 T64 G65 F66 L67 S79

F80 R81 V120 A121 D122 T123 L124 R125 Q126
K127 V128 L264 I265

lip8.1 L12 F13 A14 Q15 Y16 S17 A18 A19 A20
I34 V44 A49 T50 F51 L52 F66 L67 A68 L69
D70 N71 T72 N73 K74 L75 I76 V77 S79
H135 P136 D137 Y138 R139 V140 V141 T143

lip8.2 L12 F13 A14 Q15 Y16 S17 A18 A19 A20
I34 V44 A49 T50 F51 L52 Y53 S54 F55 G65
F66 L67 A68 L69 D70 N73 L75 I76 V77 L78
S79 T123 L124 R125 Q126 K127 V128 E129 D130
A131 T143

lip9.0 L6 F10 N25 N26 D27 A28 A30 G31 T50
F51 L52 Y53 S54 F55 E56 G65 F66 L67 A68
L69 I76 T123 L124 R125 Q126 K127 V128 E129
D130 A131 V132 R1333 E134 H135 P136 R139 V140
V141 F142 G156 L159 R160 G161 N162 G163 Y164
D165 I166 D167 V168 F169 S170 G190 G191 T192
L193 Y194 R195 I196 Y220

lip10.0 N11 L12 Q15 Y16 I34 T35 C36 C41 P42
E43 V44 E45 K46 A47 D48 A49 D70 N71 T72
N73 K74

lip11.0 F95 L96 L97 K98 E99 I100 N101 D102
C107 R108 G109 H110 D111 F113 T114 S115 A150
T153 V154 A173 P174 R175 V176 G177 N178 R179
F181 V203 P204 R205 L206 P207 P208 R209 F211
G212 Y213 S214 H215 G240 I241 D242 A243 T244
N248

lip12.0 L96 L97 K98 E99 I100 N101 D102 C104
S105 G106 C107 R108 G109 H110 T114 S115 V176
G177 N178 A180 F181 F184

lip13.0 N11 L12 F13 A14 Q15 Y16 S17 A182 A19
A20 Y21 N26 I34 C36 A40 C41 P42 E43 V44
A49 F55 E56 V63 T64 G65 F66 L67 A68 D70
N73 L75 I76 V77 L78 S79 F80 R81 G82 S83
R84 W89 W117 L124 V128 V141 F142 T143 G144
H145 S146 L147 G148 G149 A150 L151 A152 A155

lip14.0 Q9 F10 N11 F13 A14 S17 Y21 R81 G82
S83 R84 S85 I86 E87 N88 W89 I90 G91 N92
L93 F113 T143 G144 H145 S146 L147 G149 A150
T153 V168 F169 S170 Y171 A173 P174 R175 V176
L193 Y194 R195 I196 T197 D201 V203 P204 L206
P207 H215 H258 Y261 F262 I265 G266

lip15.0 N11 L12 F13 A14 Q15 Y16 S17 A18 A19
A20 Y21 C22 G23 K24 N25 N26 D27 A28 I34

T35 C36 A40 C41 P42 E43 V44 E45 K46 A47
A49 F51 L52 Y53 S54 F55 E56 T64 G65 F66
L67 S79 F80 R81 T123 L124 K127 L264 I265

lip16.0 A14 E87 I90 H145 G172 I196 T197 H198
T199 N200 D201 I202 P204 R205 W221 I222 K223
S224 G225 T226 G246 N247 N254 I252 P253 D254
I255 P256 A257 H258 L259 W260 Y261 F262 G263
I265

lip17.0 E1 V2 F7 F10 G177 N178 R179 A180
F181 A182 E183 F184 L185 T186 L193 R195 H198
T199 G212 S214 H215 S216 S217 P218 E219 Y220
W221 I222 K223 S224 G225 T226 V228 P229 V230
T231 R232 N233 D234 I235 V236 K237 I238 E239
G240 I241 D242 A243 T244 G245 G246 I262

lip18.0 Q9 F13 Y16 T32 N33 I34 C41 P42 E43
V44 E45 K46 A47 D48 A49 T50 F51 L52 L67
A68 L69 D70 N71 T72 N73 L75 I76 V128
V132 H135 P136 D137 Y138 R139 V140 V141 F142
Y164 D165 I166 D167 F169 Y194

For Amylase, the following amino acid residues belong to the epitope area that correspond to each epitope sequence indicated in Table 5:

je1.1 N2 G3 T4 R33 P346 Y349 I352 L353 T354
R355 P360 V362 D366 Y367 M378 K379 A380 K381
I382 D383 P384 I385 L386 E387 A388 R389 Q390

N391 F392 A393 Y394 I450 T451

je1.2 Y57 D58 Y60 D61 F65 N66 Q67 L104 G105
G106 A107 D108 A109 T110 E111 A135 W136 T137
K138 F139 D140 F141 P142 G143 R144 G145 N146
T147 Y148 S149 F151 K152 W153 R154 F158

je2.1 M6 Y8 E10 W11 H12 D26 L30 R33 V325
D326 N327 H328 D329 S330 Q331 P332 G333 E334
E337 F339 K345 Y349 V362 F363 Y364 G365 D366
Y367 Y368 G369 I370 P371 T372 H373 S374 V375
P376 A377 M378 K379 I382 D383 L386

je2.2 L289 L293 V314 P318 T323 F324 V325 D326 F339
K345 P346 L347 A348 Y349 A350 L351 I352 L353
T354 R355 F356 Q357 G358 Y359 P360 S361 V362
F363 Y364 G365 D366 Y367 Y368 G369 P376 A377
M378 K379 I382 I385 R389 Q397

je2.3 N102 V116 E117 V118 P120 R123 D159 G160 V161
D162 W163 Q168 F169 Q170 N171 R172 I173 Y174
K175 A182 W183 D184 V187 D188 N193 Y194 D195
Y196 L197 M198 Y199 A200 D201 V202 H236

je2.4 T1 N2 T4 M6 Y8 D26 L30 R31 N32 R33
G34 I35 V325 D326 F339 K345 Y349 L353 V362
F363 Y364 G365 D366 Y367 Y368 G369 I370 P376
A377 M378 K379 I382 D383 P384 I385 L386 E387
A388 R389 Q390 N391 F392 Y394 H417

je3.1 M6 Q7 Y8 F9 E10 L13 H19 W20 N21 R22
L23 R24 D25 D26 A27 S28 N29 L30 R31 N32

	R33	I385	W39	I40	P41	P42	A43	W44	V52	G53
	Y54	Y75	A87	L88	N91	V93	D98	V100	Y364	
	Y368									

je3.2	Y8	F9	W11	H19	W20	W39	I40	P41	P42	A43
	W44	D51	V52	G53	Y54	G55	A56	Y75	D98	V99
	V100	M101	N102	H103	L104	D195	L197	M198	A200	
	D201	V202	R230	I231	D232	A233	V234	K235	H236	
	I237	E262	H328							

je3.3	Y8	F9	H19	W20	W39	I40	P41	P42	A43	W44
	K45	G46	T47	V52	G53	Y54	G55	A56	Y57	D58
	L59	Q67	K68	Y75	D98	V100	L104	G105	G106	
	A107	D108	A109	T110	E111	A135	W136	T137	K138	
	F139	D140	F141	P142						

je4.1	L23	D25	D26	A27	S28	N29	L30	R31	N32	R33
	G34	I35	T36	I38	A84	I85	H86	A87	L88	K89
	N90	N91	G92	V93	Q94	V95	Q390			

je4.2	A43	W44	K45	L59	Y60	D61	L62	G63	E64	F65
	V71	R72	T73	K74	Y75	G76	T77	R78	S79	Q80
	L81	E82	S83	Y148	W219	Y220	T223	L224		

Example 4

Having identified 'antibody binding peptide' sequences (e.g. "SDFGHKV") and by consensus analysis also ''epitope patterns'' (e.g. >DF>>K>), one can identify potential epitope sequences on the 3-dimensional surface of a parent protein (=acceptor protein) using the following method:

The anchor amino acid residues are transferred to a three dimensional structure of the protein of interest, by colouring D red, F white and K blue. Any surface area having all three residues within a distance of 18Å, preferably 15Å, more preferably 12Å, is then claimed to be an epitope. The relevant distance can easily be measured using e.g. molecular graphics programs like InsightII from Molecular Simulations Inc.

The residues in question should be surface exposed, meaning that the residue should be more than 20% surface exposed, preferably more than 50% surface exposed, more preferably 70% surface exposed. The percentage "surface accessible area" of an amino acid residue of the parent protein is defined as the Connolly surface (ACC value) measured using the DSSP program to the relevant protein part of the structure, divided by the residue total surface area and multiplied by 100. The DSSP program is disclosed in W. Kabsch and C. Sander, BIOPOLYMERS 22 (1983) pp. 2577-2637. The residue total surface areas of the 20 natural amino acids are tabulated in Thomas E. Creighton, PROTEINS; Structure and Molecular Principles, W.H. Freeman and Company, NY, ISBN: 0-7167-1566-X (1984).

Substitutions of one or more residue (s) within 18Å, preferably 15Å, more prefereably 12Å, around the geometrical center

of the residues involved in the epitope, for a bigger or smaller residues, may destroy the epitope, and make the protein less antigenic.

Residues involved in epitope is 2, preferably 3 and more prefereably 4

Example 5

Production, selection, and evaluation of enzyme variants with reduced antigenicity or immunogenicity.

Epitope sequences and hot-spots amino acids were mutated using standard techniques know to the person skilled in the field (e.g. site-directed mutagenesis, error-prone PCR- see for example Sambrook et al. (1989), Molecular Cloning. A Laboratory Manual, Cold Spring Harbour, NY).

In the examples shown below, variants were made by site-directed mutagenesis. Amino acid exchanges giving new epitopes or duplicating existing epitopes, according to the information collected in the epitope-database (See Example 1), were avoided in the mutagenesis process.

Enzyme variants were screened for reduced binding of antibodies raised against the backbone enzyme. Antibody binding was assessed by competitive ELISA as described in the Methods section.

Variants with reduced antibody binding capacity were further evaluated in the mouse SC animal model (See methods section).

The following variants showed reduced IgE and/or reduced IgG levels in the mouse model:

Parent protein	Mutations	Target epitope sequences	%IgG response	%IgE response
Savinase	D181N	Sav11.0; Sav15.0 and Sav18.1. Hot spot amino acid.	50	19
Savinase	R170L;Q206E	Sav9,4; Sav10,4; Sav1.1; and Sav19.2	5	34
Savinase	R170L, S57P	Sav9,4; Sav10,4	45	12
Savinase	R247E	Sav2.3, Sav6.1, Sav18.2 Hot spot amino acid.	75	30
Savinase	R247Q	Sav2.3, Sav6.1, Sav18.2 Hot spot amino acid.	17	20
Savinase	R247H	Sav2.3, Sav6.1, Sav18.2 Hot spot amino acid.	40	27
Savinase	R247K	Sav2.3, Sav6.1, Sav18.2 Hot spot amino acid.	74	34

Example 6

Production, selection, and evaluation of enzyme variants with reduced antigenicity or immunogenicity.

Hot-spots or epitopes were mutated using techniques known to the expert in the field (e.g. site-directed mutagenesis, error-prone PCR).

In the examples showed below, variants were made by site-directed mutagenesis. Amino acid exchanges giving new epitopes or duplicating existing epitopes according to the information collected in the epitope-database, were avoided in the mutagenesis process.

Enzyme variants were screened for reduced binding of antibodies raised against the backbone enzyme. This antibody binding was assessed by established assays (e.g. competitive ELISA, agglutination assay).

Variants with reduced antibody binding capacity were further evaluated in animal studies.

Mice were immunised subcutanuous weekly, for a period of 20 weeks, with 50 µl 0.9% (wt/vol) NaCl (control group), or 50 µl 0.9% (wt/vol) NaCl containing 10 µg of protein. Blood samples (100 µl) were collected from the eye one week after every second immunization. Serum was obtained by blood clothing, and centrifugation.

Specific IgG1 and IgE levels were determined using the ELISA specific for mouse or rat IgG1 or IgE. Differences between

data sets were analysed by using appropriate statistical methods.

A. Site-directed mutagenesis of amino acids defining epitopes, with an effect on IgG1 and/or IgE responses in mice.

Epitope: A172/A169 R170 A194 G193 N261
Pattern: A R > R > A > N
Antibody: IgG1 + IgE
Backbone: Savinase

The variant carried mutation R170F.
In a competitive IgE ELISA, this variant was less effective in competing for anti-savinase antibodies, giving a 15% lower endpoint inhibition as compared to the savinase backbone.

Mouse studies revealed an 80% reduction of the specific IgE levels, as compared to savinase backbone ($p<0.01$). The IgG1 levels were not significantly affected.

Epitope: S216 E219 Y220
Pattern: E Y > M
Antibody: IgG1
Backbone: Lipoprime

The variant carried mutation S216W.

In a competitive IgG ELISA, the variant was less effective in competing for Lipolase antibodies, giving a 38% decrease in endpoint inhibition as compared to the enzyme backbone.

Mouse studies revealed a 69% decrease in specific IgG1 levels, compared to the lipolase backbone ($p<0.05$). The IgE levels were not significantly affected.

B. Site-directed mutagenesis of epitopes, with examples of epitope duplication, and new epitope formation, respectively, predicted by the epitope-database.

Epitope: T143 N173 N140 E136 L135
Pattern: S/T N N > E L
Antibody: IgG1
Backbone: Savinase

The variant carried mutation E136R.
In a competitive IgG ELISA, the variants was less effective in competing for savinase antibodies, giving a 38% decrease in endpoint inhibition as compared to the savinase backbone.

Mouse studies revealed a dramatic increase in specific IgG1 levels, compared to savinase backbone ($p<0.01$). The IgE levels were not significantly affected.

Mutation E136R establishes an IgG1 epitope of the R Y P R/K pattern, previously identified on PD498. Apparently, this new epitope was more antigenic in mice than the existing epitope. The introduction of a savinase unrelated epitope on the savinase backbone could explain the observed discrepancy between competitive ELISA and animal studies.

In this example, it was found that using information derived exclusively from screening phage libraries with anti-PD498 antibodies (to identify the R Y P R/K epitope pattern of Table 2) one could predict the outcome of a genetic engineering experiment for Savinase in which the E136R mutation created the PD498-epitope on the Savinase surface, leading to increased immunogenicity of this Savinase variant. This demonstrates that the epitope patterns identified may be used to predict the effect on immunogenicity of substitutions in proteins that are different from the parent protein(s) used to identify the epitope pattern.

C. Site-directed mutagenesis of amino acids defining epitope areas, with a differential effect on IgG1 and IgE antibody levels in mice, and an inhibiting effect on IgG binding, respectively.

Epitope: A172/A169 R170 A194 G193 N261
Pattern: A R > R > A > N
Antibody: IgG1 + IgE
Backbone: Savinase
Epitope area: P131, S132, A133, L135, E136, V139, A151, A152, S153, G161, S162, I165, S166, Y167, P168, Y171, N173, A174, A176, Q191, Y192, G195, L196, R247, S259, T260, L262, Y263, G264.

The variant was different at position Y167 by the mutation Y167I.
In a competitive IgE ELISA, the variant was less effective in competing for anti-savinase antibodies, giving a 8% lower endpoint inhibition as compared to the its backbone.

Mouse studies revealed an 75% reduction of the specific IgE levels, as compared to the backbone ($p<0.01$). In contrast, the IgG1 levels were dramatically increased ($p<0.01$).

Epitope: T143 N173 N140 E136 L135
Pattern: S/T N N > E L
Antibody: IgG1
Backbone: Savinase
Epitope area: V10A, I107, A108, L111, E112, G115, S132, A133, T134, Q137, A138, V139, S141, A142, S144, R145, G146, V147, V149, Y167, P168, Y171, A172, A174, M175, N243, R247.

While variant no. 1 was mutated at the epitope position (N140D), variant no. 2 was mutated at N140 (N140D), but also at the epitope area position (A172D).

In a competitive IgG ELISA, variant no. 1 was less effective in competing for anti-savinase antibodies, as compared to savinase. This variant revealed a 21% lower endpoint inhibition as compared to the its backbone.

Variant no. 2 resulted in an endpoint inhibition that was 60% lower as compared to savinase, and 40% as compared to variant no. 1.

Example 7

Conjugation of Savinase variant E136K with activated bis-PEG-1000

4,9 mg of the Savinase variant was incubated in 50 mM Sodium Borate pH 9.5 with 12 mg of N-succinimidyl carbonate activated bis-PEG 1000 in a reaction volume of approximately 2 ml. The reaction was carried out at ambient temperature using magnetic stirring while keeping the pH within the interval 9.0-9.5 by addition of 0.5 M NaOH. The reaction time was 2 hours.
The derivatives was purified and reagent excess removed by size exclusion chromatography on a Superdex-75 column (Pharmacia) equilibrated in 50 mM Sodium Borate, 5mM Succinic Acid, 150 mM NaCl, 1 mM $CaCl_2$ pH 6.0.
The conjugate was stored at -20°C, in the above described buffer.

Compared to the parent enzyme variant, the protease activity of the conjugate was retained (97% using Dimethyl-casein as substrate at pH 9).

Example 8

Competitive ELISA was performed according to established procedures. In short, a 96 well ELISA plate was coated with the parent protein. After proper blocking and washing, the coated antigen was incubated with rabbit anti-enzyme polyclonal antiserum in the presence of various amounts of modified protein (the competitior).

The amounts of residual rabbit antiserum was detected by pig anti-rabbit immunoglobulin, horseraddish peroxidase labelled.

Epitope: T143 N173 N140 E136 L135
Pattern: S/T N N > E L
Antibody: IgG1
Backbone: Savinase

Mutation: E136K
Modification: bis-NHS-PEG1000

The data show that the derivative (60% endpoint inhibition) has reduced capacity to bind enzyme specific immunoglobulines, as compared to the parent protein (100% endpoint inhibition).

CLAIMS

1. A method of selecting a protein variant having modified immunogenicity as compared to a parent protein, comprising the steps of:

a) obtaining antibody binding peptide sequences,

b) using the sequences to localise epitope sequences on the 3-dimensional structure of the parent protein,

c) defining an epitope area including amino acids situated within 5 Å from the epitope amino acids constituting the epitope sequence,

d) changing one or more of the amino acids defining the epitope area of the parent protein by genetic engineering mutations of a DNA sequence encoding the parent protein,

e) introducing the mutated DNA sequence into a suitable host, culturing said host and expressing the protein variant, and

f) evaluating the immunogenicity of the protein variant using the parent protein as reference.

2. The method according to claim 1, wherein the sequences of step a) are obtained by screening a random peptide display package library with antibodies raised against any protein of interest and sequencing the amino acid sequence of the antibody binding peptide, or the DNA sequence encoding the antibody binding peptide.

3. The method according to claim 2, wherein antibodies for screening the random peptide display package library are raised against the parent protein.

4. The method according to claims 2-3, wherein the peptide display package library is a phage display library.

5. The method according to claims 2-4, wherein the peptides of the peptide display package library are oligopeptides having from 5 to 25 amino acids.

6. The method accoding to claim 1, wherein the antibody binding peptide sequences of step a) are obtained by screening a library of known peptides related to the primary sequence of any protein of interest, with antibodies raised against the protein of interest.

7. The method according to any of the preceding claims, wherein epitope patterns are identified by sequence alignment of antibody binding peptide sequences and these epitope patterns are used to quide localisation of epitope sequences on the 3-dimensional structure of the parent protein.

8. The method according the any of the preceding claims, wherein the epitopee area of step c) equals the epitope sequence.

9. The method according to any of the preceding claims, wherein hot spot amino acids of the parent protein are identified.

10. The method according to any of the preceding claims, wherein the epitope area is changed by substituting, adding and/or deleting at least one amino acid of the epitope area.

11. The method according to claim 10, wherein the epitope area is changed by substituting, adding and/or deleting at least one hot spot amino acid.

12. The method according to claims 10-11, wherein amino acids in the epitope area are changed by substituting and/or inserting at least one amino acid by an amino acid which render the substituted and/or inserted amino acid a target for covalent conjugation to an activated polymer.

13. the method according to claim 12, wherein the amino acid for substitution and/or insertion is selected from the group consisting of K, C, D, E.

14. The method according to claim 12, wherein the molecule for covalent conjugation is selected from the group of activated synthetic or natural polymers.

15. The method according to claim 14, wherein the activated synthetic polymer is a polyethylene glycol.

16. The method according to any of the preceding claims, wherein the immunogenicity is measured by competitive ELISA.

17. The method according to any of the preseding claims, wherein the protein variant has reduced allergenicity.

18. The method according to claim 17, wherein the allergenicity of the protein variant is below 75%, preferably below 50%, more preferably below 25% of the allergenicity of the parent protein.

19. The method according to any of the preceding claims, wherein the parent protein is an enzyme or an environmental allergen or a pharmaceutical protein.

20. The method according to claim 19, wherein the enzyme is selected from the group consisting of glycosyl hydrolases, carbohydrases, peroxidases, proteases, lipases, phytases, polysaccharide lyases, oxidoreductases, transglutaminases and glucoseisomerases.

21. The method according to claim 19, wherein the environmental allergen is selected from the group consisting of pollen, dust mites, mammals, venoms, fungi, food allergens or other plant allergens.

22. A protein variant obtainable by a method according to claims 1-21.

23. A protein variant having modified immunogenicity as compared to its parent protein, wherein the amino acid sequence of the protein variant differs from the amino acid sequence of the parent protein with respect to at least one epitope area of the parent protein.

24. A protein variant according to claims 22 or 23, wherein the epitope areas are defined on the parent protein structure by being localised less than 5 Å from any of the following

epitope patterns: P > S/T D P G; P > > D A G; > P > R D T G; P > S/T D P G; > R Y > K/R; > R S A; > G > > A G; V H > G >; A > I D P R/K; A R > A; Q > Y > D >; > P > > A P > S; R/K R F > N; D/E Q I F F T; A > > > > Y P >; L > G R S; R P P R; > E Y; > P > > P A P > S; > K L > >; K Q S; > K L > >; Y I > K L; R Q > > D/E; N > > E L.

25. The protein variant according to claims 22 or 23, wherein the epitope areas correspond to antibody binding peptide sequences reactive to antibodies raised against the parent protein.

26. The protein variant according to claims 22-25, wherein the epitope pattern is a IgE epitope pattern.

27. The protein variant according to claims 22-26, wherein at least one hot spot amino acid is substituted or deleted.

28. The protein variant according to claims 22-27, wherein the allergenicity of the protein variant is below 75%, preferable below 50%, more preferably below 25% of the allergenicity of the parent protein.

29. The protein variant according to claims 22-28, wherein the protein variant is an environmental allergen, preferable an allergen selected from the group consisting of pollen, dust mites, mammals, venoms, fungi, food allergens or other plant allergens.

30. The protein variant according to claims 22-28, wherein the protein variant is an antifungal peptide or antimicrobial peptide.

31. The protein variant according to claim 29, wherein the allergen is pollen allergen Bet v1 (SEQ ID NO. 6).

32. The protein variant according to claim 31, wherein the pollen allergen of Bet v1 comprises one or more of the following substitutions:

position P31 to A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y;
position A34 to C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y;
position P35 to A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y;
position A37 to C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y;
position S39 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y;
position S40 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y;
position P59 to A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y;
position L62 to A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y;
position P63 to A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y.

33. The protein variant according to claim 32, wherein the pollen allergen comprises one or more of the following substitutions:

position P31 to A, G, L, or S;

position A34 to D, E, F, H, K, N, P, Q, R, W, or Y;
position P35 to A, G, L, or S;
position A37 to D, E, F, H, K, N, P, Q, R, W, or Y;
position S39 to D, E, F, H, K, N, P, Q, R, W, or Y;
position S40 to D, E, F, H, K, N, P, Q, R, W, or Y;
position P59 to A, G, L, or S;
position L62 to D, E, F, H, K, N, P, Q, R, W, or Y;
position P63 to A, G, L, or S.

34. The protein variant according to claims 22-28, wherein the protein variant is an enzyme.

35. The protein variant according to claim 34, wherein the enzyme is B.lentus subtilisin, PD498, Lipolase™ (SEQ ID NO 1), Amylase SP722 (SEQ ID NO 2) or the corresponding variant JE-1, Carezyme full length (SEQ ID NO 5) or Carezyme core (SEQ ID NO 4) or Laccase (SEQ ID NO 3).

36. The protein variant according to claim 35, wherein the B. lentus subtilisin comprises one or more of the following substitutins:

Position 2 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 3 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 8 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 10 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 12 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 15 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 16 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 18 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 19 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 21 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 22 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 23 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 24 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 25 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 26 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 27 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 35 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 38 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 39 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 40 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 41 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 42 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 43 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 44 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 45 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 59 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 60 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 73 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 74 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 75 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 79 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 80 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 81 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 86 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 87 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 88 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 89 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 91 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 93 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 108 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 109 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 111 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 112 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 116 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 117 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 124 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 126 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 127 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 135 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 136 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 139 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 140 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 141 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 142 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 143 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 144 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 145 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 146 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 147 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 148 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 149 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 151 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 156 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 163 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 168 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 169 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 170 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 172 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 173 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 174 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 179 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 181 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 185 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 186 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 187 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 188 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 189 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 190 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 191 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 192 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 193 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 194 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 195 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 196 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 197 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 203 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 206 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 208 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 213 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 214 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 215 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 217 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 232 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 233 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 234 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 235 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 236 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 240 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 242 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 245 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 246 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 247 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 250 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 251 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 254 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 261 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 265 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 267 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 268 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 269 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 271 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 272 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 273 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 275 A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

37. The protein variant according to claim 36, wherein the protease comprises one or more of the following substitutions:

Position 2 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 8 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 10 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 12 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 15 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 16 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 18 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 19 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 21 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 35 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 38 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 39 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 41 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 42 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 43 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 44 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 45 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 59 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 60 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 73 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 74 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 75 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 79 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 80 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 81 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 86 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 88 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 89 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 108 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 111 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 112 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 116 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 117 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 124 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 126 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 127 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 135 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 136 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 140 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 141 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 142 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 143 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 144 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 145 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 146 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 147 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 148 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 149 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 151 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 156 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 163 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 169 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 170 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 172 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 173 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 174 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 179 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 181 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 185 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 186 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 187 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 188 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 189 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 190 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 191 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 192 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 193 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 194 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 195 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 196 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 197 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 203 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 206 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 208 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 213 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 214 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 240 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 242 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 245 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 247 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 251 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 261 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 265 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 267 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 269 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 271 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 272 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 273 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y
Position 275 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

38. The protein variant according to claim 36, wherein the protease comprises one or more of the following substitutions:

Position 21 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 24 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 41 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 79 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 81 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 112 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 145 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 173 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 174 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 181 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 247 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 251 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

Position 275 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

39. The protein variant according to claim 35, wherein PD498 comprises one or more of the following substitutions:

Position -6 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position -5 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position -4 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position -2 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position -1 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 1 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 2 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position *3a to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 6 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 7 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 8 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 10 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 12 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 13 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 14 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 15 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 16 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 17 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 18 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 19 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 21 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 22 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 23 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 24 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 25 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 28 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position *28a to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 29 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 33 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 35 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 37 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 42 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 43 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 44 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position *44a to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position *44aa to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 46 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 48 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 53 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 55 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 56 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 57 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 75 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 81 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 86 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 87 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 88 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 89 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 91 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 92 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 94 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 96 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 108 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 111 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 112 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 113 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 114 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 115 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 117 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 119 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 120 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 133 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 134 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 135 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 136 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 137 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 138 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 139 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 141 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 142 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 144 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 145 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 146 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 147 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 148 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 160 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 161 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 162 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 163 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 167 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 169 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 170 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 171 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 174 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 176 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 182 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 186 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 191 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 192 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 193 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 196 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 197 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 198 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 205 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 215 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 217 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 219 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 222 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 233 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 234 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 235 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 236 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 237 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 238 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 239 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 240 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 243 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 246 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 247 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 249 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 252 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 254 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 270 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 273 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 274 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 275 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 276 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

40. The protein variant according to claim 39, wherein the protease comprises one or more of the following substitutions:

Position 16 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 21 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 22 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 23 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 28 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 29 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 44 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 46 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 48 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 53 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 55 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 56 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 57 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 86 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 89 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 91 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 94 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 108 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 111 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 112 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 113 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 115 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 117 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 119 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 120 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 133 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 134 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 135 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 136 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 137 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 138 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 141 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 142 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 144 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 145 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 146 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 161 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 162 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 186 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 191 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 192 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 205 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 215 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 217 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 219 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 222 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 233 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 234 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 236 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 237 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 239 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 240 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 243 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 246 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 276 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

41. The protein variant according to claim 39, wherein the protease comprises one or more of the following substitutions:

Position 14 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 15 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 16 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 22 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 37 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 42 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 44 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 46 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 48 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 53 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 91 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 94 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 108 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 112 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 113 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 137 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 141 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 145 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 233 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 234 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 236 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

Position 239 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 240 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 243 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 274 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;
Position 276 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W or Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

42. The protein variant according to claim 35, wherein Lipolase™ (SEQ ID NO 1) comprises one or more of the following substitutions:

Position GLN 15 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position TYR 16 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 18 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 19 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 20 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ASN 25 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ASN 26 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLU 43 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position VAL 44 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 46 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 47 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 49 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 52 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 53 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 54 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 65 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 67 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 68 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 69 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 72 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 74 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 75 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position VAL 77 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 79 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 81 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 83 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 85 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TRP 89 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 96 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 97 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 98 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLU 99 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 106 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position CYS 107 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 108 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 123 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 124 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 127 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLU 129 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 131 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position VAL 132 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 138 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position VAL 140 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 147 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 150 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 153 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 164 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 165 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 167 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 170 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 171 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 172 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 173 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 174 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 175 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position VAL 176 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 177 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 179 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 182 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 194 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position HIS 198 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASN 200 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 207 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 208 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 209 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 212 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 214 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position HIS 215 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 216 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 217 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 218 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLU 219 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 220 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 223 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 224 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 234 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 235 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 237 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 238 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 242 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 243 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 250 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 253 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 254 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 255 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 256 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 261 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

43. The protein variant according to claim 42, wherein the lipase comprises one or more of the following substitutions:

Position GLY 65 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LEU 67 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ARG 81 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position SER 83 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position SER 85 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LEU 96 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LEU 97 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LEU 124 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position GLU 129 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position TYR 164 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ARG 179 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 182 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 207 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 208 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 209 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 212 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 214 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position HIS 215 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 216 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 217 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 218 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLU 219 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 220 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 243 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 250 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 253 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

44. The protein variant according to claim 42, wherein the lipolase comprises one or more of the following substitutions:

Position PRO 207 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 208 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position SER 214 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position SER 216 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position SER 217 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 243 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 250 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 253 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position TYR 53 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

45. The protein variant according to claim 35, wherein Amylase SP722 (SEQ ID NO 2) or variant thereof (JE-1) comprises one or more of the following substitutions:

Position TYR 8 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 25 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 26 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 28 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASN 29 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 31 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 41 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 42 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 54 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 57 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 62 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 63 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 76 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 78 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 79 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 88 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 92 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASN 102 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 107 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 108 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 138 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 140 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 142 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 144 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLN 170 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 173 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 195 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 196 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 232 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 233 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLN 331 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 349 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 352 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLN 357 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 366 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 367 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 368 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 370 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 380 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 381 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 382 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 383 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 384 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 386 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 389 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLN 390 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

46. The protein variant according to claim 45, wherein the amylase comprises one or more of the following substitutions:

Position PRO 41 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 42 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 109 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LYS 138 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ASP 140 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 142 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ARG 144 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ASP 366 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position TYR 367 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position TYR 368 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 380 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LYS 381 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 382 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ASP 383 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 384 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ARG 389 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

47. The protein variant according to claim 35, wherein Carezyme core (SEQ ID NO 4) or Carezyme full length (SEQ ID NO 5) comprises one or more of the following substitutions:

Position ALA to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ASP 2 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ARG 7 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position LYS 20 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 23 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PHE 29 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position GLN 36 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 37 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ILE 38 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 40 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PHE 41 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 44 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 54 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLN 59 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 61 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TRP 62 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 83 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 84 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 95 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 96 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 98 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 100 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 101 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 133 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 134 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 136 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 146 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 153 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 160 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 162 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LEU 163 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position LYS 164 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 165 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 168 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TRP 169 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 170 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PHE 174 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASN 176 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 177 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 178 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 180 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position SER 183 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 191 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 195 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position THR 197 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ARG 200 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

48. The protein variant according to claim 47, wherein the carhydrase comprises one or more of the following substitutions:

Position LYS 20 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 23 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position PRO 27 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position ALA 83 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position GLY 84 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 95 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 96 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 98 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 100 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 101 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

with the proviso that the amino acids of the parent enzyme are substituted to another amino acid.

49. The protein variant according to claim 35, wherein Laccase comprises one or more of the following substitutions:

Position THR 10 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 12 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 23 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 30 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 51 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 53 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 59 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 60 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 78 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 79 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 80 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 113 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 114 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 118 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 155 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 157 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 165 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 166 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 168 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 175 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 180 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 211 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ASP 213 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 234 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLN 241 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 265 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 296 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 300 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 349 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 350 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position PRO 378 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 379 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ALA 389 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 390 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position ARG 409 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position SER 410 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position TYR 416 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position GLY 430 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position THR 432 to A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;

Position	ASP	434	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position	THR	442	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position	ASP	443	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position	PRO	445	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position	GLY	446	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position	ASP	469	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y;
Position	THR	473	to	A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, Y

with the proviso that the amino acids of the parent protein is substituted to another amino acid.

50. A composition comprising a protein variant as defined in any of claims 22-49.

51. The composition according to claim 50, wherein the composition is in form of a pharmaceutical composition such as a vaccine.

52. The composition according to claim 50, wherein the compositions is in form of a industrial composition such as, a detergent composition, personal care composition.

53. The use of the composition as defined in claim 50 for the production of a pharmaceutical.

54. The use of the composition as defined in claim 50 for industrial application.

55. A DNA construct comprising a DNA sequence encoding a protein variant as defined in any of claims 22-49.

56. An expression vector comprising a DNA construct according to claim 55.

57. A host cell which is capable of expressing a polypeptide and comprising a DNA construct as defined in claim 55.

58. A host cell which is capable of expressing a polypeptide and which is transformed by an expression vector according to claim 56.

59. A host according to claims 57 or 58, which is a fungal cell, an insect cell, a mammalian cell, or a plant cell.

60. A method of producing a protein variant having reduced immunogenicity as compared to the parent protein, comprising:

- culturing a host according to any of claims 57-59 in a suitable culture medium to obtain expression and secretion of the protein into the medium, followed by

- isolation of the protein from the culture medium.

61. A kit for characterizing specificity of the allergic response of a patient, comprising a set of antibody binding peptide sequences corresponding to at least one epitope on at least one potential allergen.

62. The kit according to claim 61, for which the antibody binding sequences each are specific for one out of a known range of allergens, such that the characterization of allergic specificity becomes less susceptibility to cross-reactivity interferences.

63. A kit according to claims 61-62, which further comprises other diagnostic reagents, which facilitate determination of the serum response to each of the antibody binding sequences.

64. A kit according to claims 60-63, which further comprises allergen vaccines, which can be administered to the patient according to the test results obtained using the antibody binding sequences.

ABSTRACT

TITLE: PROTEIN VARIANT HAVING MODIFIED IMMUNOGENICITY

The present invention relates to a method of selecting a protein variant having modified immunogenicity as compared to the parent protein comprising the steps obtaining antibody binding peptide sequences, using the sequences to localise epitope sequences on the 3-dimensional structure of parent protein, defining an epitope area including amino acids situated within 5 Å from the epitope amino acids constituting the epitope sequence, changing one or more of the amino acids defining the epitope area of the parent protein by genetical engineering mutations of a DNA sequence encoding the parent protein, introducing the mutated DNA sequence into a suitable host, culturing said host and expressing the protein variant, and evaluating the immunogenicity of the protein variant using the parent protein as reference.

The invention further relates to the protein variant and use thereof, as well as to a method for producing said protein variant.

SEQUENCE LISTING

SEQ ID NO 1: Lipolase:

Met Arg Ser Ser Leu Val Leu Phe Phe Val Ser Ala Trp Thr Ala Leu
Ala Ser Pro Ile Arg Arg Glu Val Ser Gln Asp Leu Phe Asn Gln Phe
Asn Leu Phe Ala Gln Tyr Ser Ala Ala Ala Tyr Cys Gly Lys Asn Asn
Asp Ala Pro Ala Gly Thr Asn Ile Thr Cys Thr Gly Asn Ala Cys Pro
Glu Val Glu Lys Ala Asp Ala Thr Phe Leu Tyr Ser Phe Glu Asp Ser
Gly Val Gly Asp Val Thr Gly Phe Leu Ala Leu Asp Asn Thr Asn Lys
Leu Ile Val Leu Ser Phe Arg Gly Ser Arg Ser Ile Glu Asn Trp Ile
Gly Asn Leu Asn Phe Asp Leu Lys Glu Ile Asn Asp Ile Cys Ser Gly
Cys Arg Gly His Asp Gly Phe Thr Ser Ser Trp Arg Ser Val Ala Asp
Thr Leu Arg Gln Lys Val Glu Asp Ala Val Arg Glu His Pro Asp Tyr
Arg Val Val Phe Thr Gly His Ser Leu Gly Gly Ala Leu Ala Thr Val
Ala Gly Ala Asp Leu Arg Gly Asn Gly Tyr Asp Ile Asp Val Phe Ser
Tyr Gly Ala Pro Arg Val Gly Asn Arg Ala Phe Ala Glu Phe Leu Thr
Val Gln Thr Gly Gly Thr Leu Tyr Arg Ile Thr His Thr Asn Asp Ile
Val Pro Arg Leu Pro Pro Arg Glu Phe Gly Tyr Ser His Ser Ser Pro
Glu Tyr Trp Ile Lys Ser Gly Thr Leu Val Pro Val Thr Arg Asn Asp
Ile Val Lys Ile Glu Gly Ile Asp Ala Thr Gly Gly Asn Asn Gln Pro
Asn Ile Pro Asp Ile Pro Ala His Leu Trp Tyr Phe Gly Leu Ile Gly
Thr Cys Leu *

SEQ ID NO 2: SP722

His His Asn Gly Thr Asn Gly Thr Met Met Gln Tyr Phe Glu Trp His
Leu Pro Asn Asp Gly Asn His Trp Asn Arg Leu Arg Asp Asp Ala Ser
Asn Leu Arg Asn Arg Gly Ile Thr Ala Ile Trp Ile Pro Pro Ala Trp
Lys Gly Thr Ser Gln Asn Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr
Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly

Thr Arg Ser Gln Leu Glu Ser Ala Ile His Ala Leu Lys Asn Asn Gly

Val Gln Val Tyr Gly Asp Val Val Met Asn His Lys Gly Gly Ala Asp

Ala Thr Glu Asn Val Leu Ala Val Glu Val Asn Pro Asn Asn Arg Asn

Gln Glu Ile Ser Gly Asp Tyr Thr Ile Glu Ala Trp Thr Lys Phe Asp

Phe Pro Gly Arg Gly Asn Thr Tyr Ser Asp Phe Lys Trp Arg Trp Tyr

His Phe Asp Gly Val Asp Trp Asp Gln Ser Arg Gln Phe Gln Asn Arg

Ile Tyr Lys Phe Arg Gly Asp Gly Lys Ala Trp Asp Trp Glu Val Asp

Ser Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Val Asp Met

Asp His Pro Glu Val Val Asn Glu Leu Arg Arg Trp Gly Glu Trp Tyr

Thr Asn Thr Leu Asn Leu Asp Gly Phe Arg Ile Asp Ala Val Lys His

Ile Lys Tyr Ser Phe Thr Arg Asp Trp Leu Thr His Val Arg Asn Ala

Thr Gly Lys Glu Met Phe Ala Val Ala Glu Phe Trp Lys Asn Asp Leu

Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Trp Asn His Ser Val

Phe Asp Val Pro Leu His Tyr Asn Leu Tyr Asn Ala Ser Asn Ser Gly

Gly Asn Tyr Asp Met Ala Lys Leu Leu Asn Gly Thr Val Val Gln Lys

His Pro Met His Ala Val Thr Phe Val Asp Asn His Asp Ser Gln Pro

Gly Glu Ser Leu Glu Ser Phe Val Gln Glu Trp Phe Lys Pro Leu Ala

Tyr Ala Leu Ile Leu Thr Arg Glu Gln Gly Tyr Pro Ser Val Phe Tyr

Gly Asp Tyr Tyr Gly Ile Pro Thr His Ser Val Pro Ala Met Lys Ala

Lys Ile Asp Pro Ile Leu Glu Ala Arg Gln Asn Phe Ala Tyr Gly Thr

Gln His Asp Tyr Phe Asp His His Asn Ile Ile Gly Trp Thr Arg Glu

Gly Asn Thr Thr His Pro Asn Ser Gly Leu Ala Thr Ile Met Ser Asp

Gly Pro Gly Gly Glu Lys Trp Met Tyr Val Gly Gln Asn Lys Ala Gly

Gln Val Trp His Asp Ile Thr Gly Asn Lys Pro Gly Thr Val Thr Ile

Asn Ala Asp Gly Trp Ala Asn Phe Ser Val Asn Gly Gly Ser Val Ser

Ile Trp Val Lys Arg

SEQ ID NO 3: Laccase:

GLN ILE VAL ASN SER VAL ASP THR MET THR LEU THR ASN
ALA ASN VAL SER PRO ASP GLY PHE THR ARG ALA GLY ILE
LEU VAL ASN GLY VAL HIS GLY PRO LEU ILE ARG GLY GLY
LYS ASN ASP ASN PHE GLU LEU ASN VAL VAL ASN ASP LEU
ASP ASN PRO THR MET LEU ARG PRO THR SER ILE HIS TRP
HIS GLY LEU PHE GLN ARG GLY THR ASN TRP ALA ASN GLY
ALA ASP GLY VAL ASN GLN CYS PRO ILE SER PRO GLY HIS
ALA PHE LEU TYR LYS PHE THR PRO ALA GLY HIS ALA GLY
THR PHE TRP TYR HIS SER HIS PHE GLY THR GLN TYR CYS
ASP GLY LEU ARG GLY PRO MET VAL ILE TYR ASP ASP ASN
ASP PRO HIS ALA ALA LEU TYR ASP GLU ASP ASP GLU ASN
THR ILE ILE THR LEU ALA ASP TRP TYR HIS ILE PRO ALA
PRO SER ILE GLN GLY ALA ALA GLN PRO ASP ALA THR LEU
ILE ASN GLY LYS GLY ARG TYR VAL GLY GLY PRO ALA ALA
GLU LEU SER ILE VAL ASN VAL GLU GLN GLY LYS LYS TYR
ARG MET ARG LEU ILE SER LEU SER CYS ASP PRO ASN TRP
GLN PHE SER ILE ASP GLY HIS GLU LEU THR ILE ILE GLU
VAL ASP GLY ASN LEU THR GLU PRO HIS THR VAL ASP ARG
LEU GLN ILE PHE THR GLY GLN ARG TYR SER PHE VAL LEU
ASP ALA ASN GLN PRO VAL ASP ASN TYR TRP ILE ARG ALA
GLN PRO ASN LYS GLY ARG ASN GLY LEU ALA GLY THR PHE
ALA ASN GLY VAL ASN SER ALA ILE LEU ARG TYR ALA GLY
ALA ALA ASN ALA ASP PRO THR THR SER ALA ASN PRO ASN
PRO ALA GLN LEU ASN GLU ALA ASP LEU HIS ALA LEU ILE
ASP PRO ALA ALA PRO GLY ILE PRO THR PRO GLY ALA ALA

ASN VAL ASN LEU ARG PHE GLN LEU GLY PHE SER GLY GLY
ARG PHE THR ILE ASN GLY THR ALA TYR GLU SER PRO SER
VAL PRO THR LEU LEU GLN ILE MET SER GLY ALA GLN SER
ALA ASN ASP LEU LEU PRO ALA GLY SER VAL TYR GLU LEU
PRO ARG ASN GLN VAL VAL GLU LEU VAL VAL PRO ALA GLY
VAL LEU GLY GLY PRO HIS PRO PHE HIS LEU HIS GLY HIS
ALA PHE SER VAL VAL ARG SER ALA GLY SER SER THR TYR
ASN PHE VAL ASN PRO VAL LYS ARG ASP VAL VAL SER LEU
GLY VAL THR GLY ASP GLU VAL THR ILE ARG PHE VAL THR
ASP ASN PRO GLY PRO TRP PHE PHE HIS CYS HIS ILE GLU
PHE HIS LEU MET ASN GLY LEU ALA ILE VAL PHE ALA GLU
ASP MET ALA ASN THR VAL ASP ALA ASN ASN PRO PRO VAL
GLU TRP ALA GLN LEU CYS GLU ILE TYR ASP ASP LEU PRO
PRO GLU ALA THR SER ILE GLN THR VAL VAL

SEQ ID NO 4: Carezyme Core (SwissProt accession number P43316):

SQ SEQUENCE 213 AA;
ADGRSTRYWD CCKPSCGWAK KAPVNQPVFS CNANFQRITD FDAKSGCEPG GVAYSCADQT
PWAVNDDFAL GFAATSIAGS NEAGWCCACY ELTFTSGPVA GKKMVVQSTS TGGDLGSNHF
DLNIPGGGVG IFDGCTPQFG GLPGQRYGGI SSRNECDRFP DALKPGCYWR FDWFKNADNP
SFSFRQVQCP AELVARTGCR RNDDGNFPAV QIP

SEQ ID NO 5: Carezyme full length (SwissProt accession number R15272)

SQ Sequence 305 AA;
mrsspllpsa vvaalpvlal aadgrstryw dcckpscgwa kkapvnqpvf scnanfqrit
dfdaksgcep ggvayscadq tpwavnddfa lgfaatsiag sneagwccac yeltftsgpv
agkkmvvqst stggdlgsnh fdlnipgggv gifdgctpqf gglpgqrygg issrnecdrf
pdalkpgcyw rfdwfknadn psfsfrqvqc paelvartgc rrnddgnfpa vqipssstss
pvnqptstst tststtsspp vqpttpsgct aerwaqcggn gwsgcttcva gstctkindw
yhqcl

SEQ ID NO 6: Bet v1 sequence SwissProt accession number P15494):
SQ SEQUENCE 159 AA;
GVFNYETETT SVIPAARLFK AFILDGDNLF PKVAPQAISS VENIEGNGGP GTIKKISFPE
GFPFKYVKDR VDEVDHTNFK YNYSVIEGGP IGDTLEKISN EIKIVATPDG GSILKISNKY
HTKGDHEVKA EQVKASKEMG ETLLRAVESY LLAHSDAYN